

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 19, 2005, 23:00:56 ; Search time 651 Seconds

(without alignments)

2761.710 Million cell updates/sec

Title: US-10-659-004-104

Perfect score: 1599

Sequence: 1 MARKSLVILLTPMLSTVNP.....FGIMLISIRBKIVRINF 302

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 554816 seqs, 297611598 residues

Total number of hits satisfying chosen parameters: 11089632

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=x1h
-Q=/cgn2_1/USPTO.spool/US10659004/runat 14032005 181236 15793/app query.fasta_1.455
-DB=Published Applications NA -OPM=fastapp -SUFFIX=rnpb -MINMATCH=0.1
-LOOBCU=0 -LOOEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pico -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MARLEN=200000000 -USER=US10659004@cgn 1.1 480 @runat 14032005 181236 15793
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA: *
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq: *
2: /cgn2_6/ptodata/2/pubpna/PTCT_NEW_PUB.seq: *
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq: *
4: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq: *
5: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq: *
6: /cgn2_6/ptodata/2/pubpna/PTCTUS_PUBCOMB.seq: *
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq: *
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq: *
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq: *
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq: *
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq: *
12: /cgn2_6/ptodata/2/pubpna/US10A_NEW_PUB.seq: *
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq: *
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq: *
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq: *
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq: *
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq: *
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq: *
19: /cgn2_6/ptodata/2/pubpna/US10G_NEW_PUB.seq: *
20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq: *
21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq: *
22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1599	100.0	1101	18	US-10-659-004-103 Sequence 103, App
2	1359.5	85.0	2250	19	US-10-764-604-28 Sequence 28, App1
3	1359.5	85.0	2744	15	US-10-171-581-170 Sequence 170, App
4	1359.5	85.0	2744	15	US-10-177-293-263 Sequence 263, App
5	1359.5	85.0	2744	17	US-10-172-118-1400 Sequence 1400, App
6	1359.5	85.0	2744	17	US-10-342-687-1400 Sequence 1400, App
7	1359.5	85.0	2744	18	US-10-769-612-1 Sequence 1, App1
8	1359.5	85.0	3523	14	US-10-357-930-25736 Sequence 25736, A
9	1359.5	85.0	3523	18	US-10-198-846-10007 Sequence 10007, A
10	1356.5	84.8	2268	17	US-10-188-832-176 Sequence 176, App
11	1356.5	84.8	2268	19	US-10-659-004-105 Sequence 105, App
12	1356.5	84.8	3461	17	US-10-295-027-55 Sequence 55, App1
13	1356.5	84.8	3461	17	US-10-295-027-838 Sequence 838, App
14	1356.5	84.8	3461	17	US-10-295-027-878 Sequence 878, App
15	1356.5	84.8	3461	17	US-10-058-270A-29 Sequence 29, App1
16	1356.5	84.8	3609	19	US-10-855-588-21 Sequence 21, App1
17	1355	84.7	1186	19	US-10-659-004-101 Sequence 101, App
18	1352.5	84.6	3445	17	US-10-099-322-25 Sequence 25, App1
19	1352.5	84.6	3445	17	US-10-044-564-25 Sequence 25, App1
20	1352.5	84.6	3445	19	US-10-659-004-97 Sequence 97, App1
21	1351.5	84.5	2310	19	US-10-659-004-107 Sequence 107, App
22	1349.5	84.4	2776	17	US-10-182-033-3 Sequence 3, App1
23	1347.5	84.3	3445	19	US-10-659-004-113 Sequence 113, App
24	1345.5	84.1	3445	19	US-10-659-004-111 Sequence 111, App
25	1324.5	82.8	2265	18	US-10-769-612-13 Sequence 13, App1
26	1311.5	82.0	3461	17	US-10-182-033-1 Sequence 1, App1
27	1311.5	82.0	3586	14	US-10-176-847-51 Sequence 51, App1
28	1285.5	79.1	2211	19	US-10-659-004-109 Sequence 109, App
29	1007	63.0	2811	19	US-10-855-588-23 Sequence 23, App1
30	875	54.7	1597	17	US-10-131-110-12 Sequence 12, App1
31	831	52.0	551	17	US-10-182-033-12 Sequence 12, App1
32	718	44.9	3309	18	US-10-357-930-23267 Sequence 23267, A
33	718	44.9	3309	18	US-10-357-930-29142 Sequence 29142, A
34	718	44.9	4417	10	US-09-814-353-22020 Sequence 22020, A
35	718	44.9	5231	16	US-10-269-909-39 Sequence 39, App1
36	718	44.9	5231	16	US-10-723-660-4160 Sequence 4160, App
37	718	44.9	5330	9	US-09-789-561-14 Sequence 14, App1
38	718	44.9	5330	19	US-10-883-936-14 Sequence 14, App1
39	718	44.9	5446	18	US-10-322-696-5 Sequence 5, App1
40	712	44.5	5299	18	US-10-723-860-8029 Sequence 8029, App
41	712	44.5	5748	18	US-10-322-696-2 Sequence 2, App1
42	578	36.1	912	19	US-10-659-004-99 Sequence 99, App1
43	514.5	32.2	366	10	US-09-918-995-30283 Sequence 30283, A
44	439	27.5	4573	9	US-09-880-107-1649 Sequence 1649, App
45	439	27.5	4573	10	US-09-960-706-489 Sequence 489, App

ALIGNMENTS

RESULT 1
US-10-659-004-103
; Sequence 103, Application US/10659004
; Publication No. US20050048507A1
; GENERAL INFORMATION:
; APPLICANT: Zhong et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD.
; FILE REFERENCE: 21402-608
; CURRENT APPLICATION NUMBER: US/10/659, 004
; PRIOR FILING DATE: 2003-09-09
; PRIOR APPLICATION NUMBER: 60/295, 607
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/295, 661
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/296, 404
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/296, 418
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/297, 414
; PRIOR FILING DATE: 2001-06-11

CURRENT APPLICATION NUMBER: US/10/171,581
CURRENT FILING DATE: 2002-06-14
PRIORITY APPLICATION NUMBER: 60/298,914
PRIORITY FILING DATE: 2001-06-18
NUMBER OF SEQ ID NOS: 366
SEQ ID NO 170
LENGTH: 2744
TYPE: DNA
ORGANISM: Homo sapiens
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: U41060
DATABASE ENTRY DATE: 2001-06-18
US-10-171-581-170

Alignment Scores:
Pred. No.: 1,31e-150 Length: 2744
Score: 1359.50 Matches: 301
Percent Similarity: 40.32% Conservative: 1
Best Local Similarity: 40.19% Mismatches: 0
Query Match: 85.02% Indels: 447
DB: 15 Gaps: 1

US-10-659-004-104 (1-302) x US-10-171-581-170 (1-2744)

QY 1 MetAlaArgLYbLeuSerValIleuThrPheAlaLeuSerValThrAsnPro 20
DB 138 ATGGGAGAGAGTATCTGTATCTTGATCTGACCTTGCCCTCTCTGTCAACAATCCC 197

QY 21 LeuHISGLuLeuLYaAlaAlaIlePheProGlnThrGluLYIleSerProAsnTrp 40
DB 198 CTTCTAGAACTAAAGCAGCTGCTTTCCCGCAGACCACTGAGAAATTAAGTCGAAATGG 257

QY 41 GluSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTrpHisLeuGlnGluLeu 60
DB 258 GAATCTGCATTAATAGTTGACTTGGCAATTTCCACACGCGCAATATCATCTACACAGCTT 317

QY 61 PheTYArgTYRGLyGluAsnAsnSerLeuSerValGluLYPheArgLYbLeuGln 80
DB 318 TTCTACCCGCTATGAGAAATATATCTTTGTGAGTTGAAAGGTTTGAATAATTACTTCA 377

QY 81 AsnIleGlyIleAspLYeIleLYaArgIleHisIleHisAspHisAspHisSer 100
DB 378 AATATAGGCAATGATTAAGTTAAAGATTCATATACCAATGACCAACGACCATCATCA 437

QY 101 AspHisGluHisIleSerAspHisGluLYArgHisSerAspHisGluHisIleSerAsp 119
DB 438 GACCAAGCAGCATCACTCAGACCATGAGGCTCACTCAGACCATGAGCATCACTCAGAC 497

QY 119 ----- 119

DB 498 GAGCATCACTGACCATATATCATGCTGCTTGTGTAATAATAGCGAAAGCTTTTGC 557

QY 119 ----- 119

DB 558 CCAGACATGACTCAGATAGTTCAGGTAAAGATCTAGAAACAGCGAGGAAAGAGCT 617

QY 119 ----- 119

DB 618 CACCGACAGAACATGCGCAGGTGTAAGAAATGTCAAGACAGTGTAGTGTAGTAA 677

QY 119 ----- 119

DB 678 GTGACCTCAACTGTGTACAACACTGTCTGTGAAGAACTCACTTTCTAGACAAATAGAG 737

QY 119 ----- 119

DB 728 ACTCCAGAGCTGAAAACCTTTCCCAAAAGATGTAAGCAGCTCCACCTCAACCCAGTGC 797

QY 119 ----- 119

DB 798 ACATCAAGAGCGGGGTGAGCGGCTGCTGTAGAAAACAAATGAATCTGTGAGTGAAG 857

QY 119 ----- 119

DB 858 CCCCCAAAAGCTTTATGATATCCAGAAACAAATGAATAATCTCAGAGGTTCAT 917

QY 119 ----- 119

DB 918 GCATCAAGCTACTGACATCTCATGCGATGGCATCCAGGTTCCGTGAATGCAACAGAG 977

QY 119 ----- 119

DB 978 TTCAACTATCTGTGTCCAGCCATCATCAACCAATTAATGATGATCTGTGTATTCAT 1037

QY 119 ----- 119

DB 1038 ACAAGTGAAGAAAGCTGAATTCCTCCAAAGACCTATTCATTACAATAGCCTGGGT 1097

QY 119 ----- 119

DB 1098 GGTGTTTATATAGCATTTTCATCATCATGATTTCTGTCTGTCTGTGGGGTTATCTTAGTG 1157

QY 119 ----- 119

DB 1158 CTTCTCATGAATCGGGTGTCTTTCAAATTTCTCTGAGTTCTTGTGGCACTGGCGGTT 1217

QY 119 ----- 119

DB 1218 GGGACTTGTAGTGTGATGCTTTTACACCTTCTTCACATTCATGCAAGTCACAC 1277

QY 119 ----- 119

DB 1278 CATAGTCATAGCCATGAAGAACACGACATGGAATGAAGAGACCACTTTTCAGTCAT 1337

QY 119 ----- 119

DB 1338 CTGTCTTCTCAAAACATAGAAGAAAGCTTATTTGATTCACAGTGAAGGGTCTAACA 1397

QY 119 ----- 119

DB 1398 GCTCTAGAGGCGCTATTTTCATGTTCTTGTGAACATGTCCTCATGTGATCAACAA 1457

QY 119 ----- 119

DB 1458 TTTAAGATTAAGAAAGAAAGATCAGAGAAACCTGAAATGATGATGTGAGATT 1517

QY 119 ----- 119

DB 1518 AAGAGCAGTTGTCCAAATGATGAATCTCACTTTCAACAAATGAGAGAAAGTAGATCA 1577

QY 119 ----- 119

DB 1578 GATGATGAAGTGAAGGCTATTTACAGACAGACTCACAAGAGCCCTCCACTTGTATCT 1637

QY 119 ----- 119

DB 1638 CAGCAGCTGCACTTTGGAAGAAAGAGTCAATGATGCTCATGCTCATCAACAGAA 1697

QY 119 ----- 119

DB 1698 GTTCAATGAATATGTACCCAGAGGCTCAAGATAAATGCCATTCACTTCCAGAT 1757

QY 119 ----- 119

DB 1758 ACACCTGGCAGTGAAGCATCTCATTCACACCATCATGATCATATATCTTCAT 1817

QY 120 ----- 133

DB 1818 CATCACCAACCAAAACCAACCATCTCAGCATGACAGCCGCTACTCTTCGGAGAGAG 1877

QY 134 LeuLYAspAlaGlyValAlaThrLeuAlaTrpMetValIleMetGlyAspGlyLeuHis 153

DB 1878 CTGAAGATGCGCGCTGCGCACTTTGCTGTGATGATGATGATGATGATGATGATGATGAT 1937

QY 154 AsnPheSerAspGlyLeuAlaIleGlyAlaAlaIlePheThrGluGlyLeuSerSerGlyLeu 173

```

Db      1938 AATTGACGATGCGCTAGCAATGCTGCTTTTACTGAAAGCTTATCAAGTGTTA 1997
Qy      174 SerThrSerValAlaValPheCyeHisGluLeuProHisGluLeuGlyAspPheAlaVal 193
Db      1998 AGTACTTGTGTGCTGTGTCTGTGTCATGATGCTGCTCAATGATGAGTGTGCTGTGT 2057
Qy      194 LeuLeuValAlaGlyMetThrValLysGlnAlaValLeuTyrAsnAlaLeuSerAlaMet 213
Db      2058 CTACTAAGGCTGCGATACCTTAAAGCAGCTGCTCTTATTAATGATGCTGACCCATG 2117
Qy      214 LeuAlaTyrLeuGlyMetAlaThrGlyIlePheIleGlyHisTyrAlaGluAsnValSer 233
Db      2118 CTGGCGTATCTTGAATGCAACAGAAATTTTCATTTGATGCTATGCGAAATGTTCTT 2177
Qy      234 MetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisValAlaLeuValAspMetVal 253
Db      2178 ATGTGGAATATTGCACTTACTGCTGCTTATTCATGATGATGCTGCTGATGATGCTGA 2237
Qy      254 ProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSerArgTrpGlyTyrPhePhe 273
Db      2238 CTGAAATGCTGCAACATGATGCTAGTACCATGATGATGATGCGCTGCGGATTTCTT 2297
Qy      274 LeuGlnAsnAlaGlyMetLeuLeuGlyIlePheGlyIleMetLeuLeuIleSerIlePheGlu 293
Db      2298 TTACGATGCTGCGGATGCTTTTGGGTTTGGAAATTAATTAATTTCCATATTGAA 2357
Qy      294 HisLysIleValPheArgIleAsnPhe 302
Db      2358 CATAAATCGTGTTCGTATAAATTC 2384

RESULT 4
US-10-177-293-263
; Sequence 263, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glact, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Ganavarpur, Manjula
; APPLICANT: Kamatkar, Shubhangit
; APPLICANT: Merrens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongzao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Baat Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Puzstai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 263

```

```

; LENGTH: 2744
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-177-293-263

Alignment Scores:
Pred. No.: 1,31e-150 Length: 2744
Score: 1359.50 Matches: 301
Percent Similarity: 40.32% Conservative: 1
Best Local Similarity: 40.19% Mismatches: 0
Query Match: 85.02% Indels: 447
DB: 15 Gaps: 1

US-10-659-004-104 (1-302) x US-10-177-293-263 (1-2744)
Qy      1 MetAlaArgLysLeuSerValIleLeuIleLeuThrPheAlaLeuSerValThrAsnPro 20
Db      138 ATGGCGAAGAAATATCTGTAATCTTGAATCCTGACCTTGGCCCTCTGTGCACAAATCCC 197
Qy      21 LeuHisGluLeuValAlaAlaPheProGlnThrTrpGluLysIleSerProAsnTrp 40
Db      198 CTTCAATGAATPAAAGCAGCTGCTTCCCAAGACCACTGAGAAATTAATGCCAATTGG 257
Qy      41 GluSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTyrHisLeuGlnGlnLeu 60
Db      258 GAATCTGCATTAATGTTGACTTGCACTTGGCAATTTCCACAGCGCATATCATCTACAAACGCTT 317
Qy      61 PheTyrArgTyrGlyGluAsnAsnSerLeuSerValGluGlyPheArgLysLeuGln 80
Db      318 TTCTACCGCTAATGAGAAATATATTCTTTGTTCAGTTGAAGGGTTCAGAAATTAATCTTCA 377
Qy      81 AsnIleGlyIleAspLysIleLysArgIleHisIleHisIleAspHisIleSer 100
Db      378 AATATAGCATATGATTAAGATTAAAGATTCATATACCATGACAGACCATCATCTCA 437
Qy      101 AspHisGluHisIleSerAspHisGluArgHisIleSerAspHisIleSerAsp 119
Db      438 GATCAGCATCTACTCTGACCATATGAGCGCTCATCTAGCATCATCTACAGACCAAC 497
Qy      119 ----- 119
Db      498 GAGCATCACTGACCATATCATGCTGCTTCTGTGTAATAAATAAGCAAAAGCTTTGCG 557
Qy      119 ----- 119
Db      558 CCAGACATGACTCATGATAGTTCAAGTAAGATCTTAGAAAACGCCAGGGAAAGAGCT 617
Qy      119 ----- 119
Db      618 CACGACCAAGAACATGCCAGTGTAGAAAGAAATGCAAGACAGTTGATGCTTAGTGA 677
Qy      119 ----- 119
Db      678 GTGACCTGAATGTGTACAACACTGTCTGTGAAGAACTCACTTTCTAGAGCAATAGAG 737
Qy      119 ----- 119
Db      738 ACTCCAGACCTGGAATACTTCCCAAAAGATGTAAGACAGCTCCATCCACCAAGTTC 797
Qy      119 ----- 119
Db      798 ACATCAAAAGACCGGGTGAGCCGCTGCTGTGTAGAAAACAATAATGAATCTGTGAGTAG 857
Qy      119 ----- 119
Db      858 CCCGAAAGAGCTTTATGTATTTCCAGAAACACAATGAATAATCTCAGAGAGTTTCAAT 917
Qy      119 ----- 119
Db      918 GCATCAAAAGTACTGACATCTCATGAGCATGGGCAATCCAGGTTCCGCTGATGACACAGAG 977
Qy      119 ----- 119

```

Db 978 TTCACTATCTCTGTCAGCCATCATCAACCAATGTAGTGTCTTGTCTGATTCAT 1037
 QY 119 ----- 119
 Db 1038 ACAAGTAAAAAGGCTGAATCCCTCCAAAAGCTTATTCATAAATAGCCCTGGTT 1097
 QY 119 ----- 119
 Db 1098 GGTGTTTTTAAGCATTTCATCATCAGTTTCCTGCTCTGCTGGGGTATATCTTAGTG 1157
 QY 119 ----- 119
 Db 1158 COTTCATGAATCGGGTGTCTTTTCAATTTCTCTGAGTTTCCTTGCGACCTGGCCGTT 1217
 QY 119 ----- 119
 Db 1218 GGGACTTGTAGTGTGATGCTTTTTCACCTTCCTTCACATTCATGCAAGTCACGAC 1277
 QY 119 ----- 119
 Db 1278 CATAGTCATAGCCATGAAGAACCCAGCAATGCAATGAAAAAGAGACCACTTTTCAGTCAT 1337
 QY 119 ----- 119
 Db 1338 CTGTCTTCTCAAAACATGAAGAAGTGCCTATTTTGATTCACGTGAAGGCTCAACA 1397
 QY 119 ----- 119
 Db 1398 GCTCTAGGAGGCTGTATTTTCATGTTCTGTGTGAACATGTCCTGCATTCATGATCAACAA 1457
 QY 119 ----- 119
 Db 1458 TTTAAAGATAGAGAAAAAGATCATGAAGAAACCTGAAATGATGATGATGTGAGATT 1517
 QY 119 ----- 119
 Db 1518 AAGAAGAGTGTGTCAAGTATGATCTCACTTTCAACAAATGAGAGAAAGTAGATACA 1577
 QY 119 ----- 119
 Db 1578 GATGATCGAAGTGAAGGCTATTTTACGAGACATCAAGAGCCCTCCCATTTGATTCCT 1637
 QY 119 ----- 119
 Db 1638 CAGCAGCTGCAGTCTTGGAAGAAGAAGGTCATGATGATGATCTCATCCACAGAA 1697
 QY 119 ----- 119
 Db 1698 GTCTCAATGATATGTACCCAGAGGTCAGAAATTAATGCAATTCACATTTCCAGCAT 1757
 QY 119 ----- 119
 Db 1758 ACACTCGGCGAGTCAGACGATCTCATCCACCATCATGACTACCATCATATTTCCAT 1817
 QY 120 ----- 133
 Db 1818 CATCACCAACCAAAAACCACTCTCTCACATCAGCCAGCCGCTACTCTCTGGAGAGAG 1877
 QY 134 LeuLysAspAlaGlyValAlaThrLeuAlaTrpMetValIleMetGlyAspGlyLeuHis 153
 Db 1878 CTGAAGAAGATGCGCGCGCTGCCACTTTGGCTGTGATGATGAATGAGGTGATGGCTGCAC 1937
 QY 154 AenPheSerAspGlyLeuAlaIleGlyValAlaPheThrGluGlyLeuSerSerGlyLeu 173
 Db 1938 AATTTCAGGATGCGCTACCAATGTGTCTTTTACGAAGGCTTACCAAGTGGTTTA 1997
 QY 174 SerThrSerValAlaValPheCysHisGluLeuProHisGlyLeuGlyAspPheAlaVal 193
 Db 1998 ACTACTCTCTGCTGTGTGTCTGTCTCATGCTTGCCTCATGAATTAGTGACCTTTGTCTGT 2057
 QY 194 LeuLeuLysAlaGlyMetThrValIleGlyGlnAlaValLeuTyrAsnAlaLeuSerAlaMet 213
 Db 2058 CTACTAAGGCTGGCATGACCGTTAAGCAGGCTGTCTTTATATATGATTCATGTCAGGCATG 2117

QY 214 LeuAlaTyrLeuGlyMetAlaThrGlyIlePheIleGlyHisTyrAlaGluAsnValSer 233
 Db 2118 CTGGCGATCTTGGAATGGACACAGAAATTTTCATTGCTCATTTGCTGAAAAATCTTCT 2177
 QY 234 MetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisValAlaLeuValAspMetVal 253
 Db 2178 ATGTGATATTGCACTTACCTGCTGCTGCTTATTCATGATGATGCTGCTGCTGATATGGTA 2237
 QY 254 ProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSerArgTrpGlyTyrPhePhe 273
 Db 2238 CTTGAATGCTGCACCAATGATGCTAGTACCATGATGATGACCCCTGGGGGTATTTCTT 2297
 QY 274 LeuGlnAsnAlaGlyMetLeuGlyPheGlyIleMetLeuLeuIleSerIlePheGly 293
 Db 2298 TTACGAATGCTGGATGCTTTTGGCTTTGGATTTATGTTACTTATTTCCATATTGAA 2357
 QY 294 HisLysIleValPheArgIleAsnPhe 302
 Db 2358 CATAAATCGTGTTCGTATTAATTTTC 2384
 RESULT 5
 US-10-172-118-1400
 ; Sequence 1400, Application US/10172118
 ; Publication No. US2003024374A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dai, Hongyue
 ; APPLICANT: He, Yudong
 ; APPLICANT: Linsley, Peter
 ; APPLICANT: Mao, Mao
 ; APPLICANT: Robertes, Chris
 ; APPLICANT: Van 't Veer, Laura
 ; APPLICANT: Van de Vijver, Marc
 ; APPLICANT: Bernards, Rene
 ; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
 ; FILE REFERENCE: 9301-175-999
 ; CURRENT APPLICATION NUMBER: US/10/172,118
 ; CURRENT FILING DATE: 2002-06-14
 ; PRIOR APPLICATION NUMBER: 60/380,770
 ; PRIOR FILING DATE: 2002-05-14
 ; NUMBER OF SEQ ID NOS: 2699
 ; SEQ ID NO 1400
 ; LENGTH: 2744
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; PUBLICATION INFORMATION:
 ; DATABASE ACCESSION NUMBER: NM_012319
 ; DATABASE ENTRY DATE: 2001-06-18
 US-10-172-118-1400
 Alignment Scores:
 Pred. No.: 1,336-150 Length: 2744
 Score: 1359.50 Matches: 301
 Percent Similarity: 40.32% Conserved: 1
 Best Local Similarity: 40.19% Mismatches: 0
 Query Match: 85.02% Indels: 447
 DB: 17 Gaps: 1
 US-10-659-004-104 (1-302) x US-10-172-118-1400 (1-2744)
 QY 1 MetAlaArgLysLeuSerValIleLeuIleLeuThrPheAlaLeuSerValThrAsnPro 20
 Db 138 ATGGGAGGAAGTATATCTGATCTTGATCTGACCTTGGCCCTCTGTGCAAAATCCC 197
 QY 21 LeuHisGluLeuLysAlaAlaPheProGlnThrThrGluLysIleSerProAsnTrp 40
 Db 198 CTTCAATGAACATAAAGCAGCTGCTTTCCCGACAGACCACTGAGAAAATTAAGTCCGAATTGG 257
 QY 41 GluSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTyrHisLeuGlnGlnLeu 60
 Db 258 GAATTCGCAATTAATGTTGACTTGCAATTTTCACAGGCAATATCATCTACACAGCTT 317
 QY 61 PheTyrArgTyrGlyGluAsnAsnSerLeuSerValGluGlyPheArgLysLeuGln 80

APPLICANT: Linsley, Peter S.
 APPLICANT: Mao, Mao
 APPLICANT: Roberts, Christopher J.
 APPLICANT: Van 't Veer, Laura Johanna
 APPLICANT: Van de Vijver, Marc J.
 APPLICANT: Bernards, Rene
 TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
 FILE REFERENCE: 9301-188-999
 CURRENT APPLICATION NUMBER: US/10/342,887
 CURRENT FILING DATE: 2003-01-15
 PRIOR APPLICATION NUMBER: 60/298,918
 PRIOR FILING DATE: 2001-06-18
 PRIOR APPLICATION NUMBER: 60/380,710
 PRIOR FILING DATE: 2002-05-14
 PRIOR APPLICATION NUMBER: 10/172,118
 PRIOR FILING DATE: 2002-06-14
 NUMBER OF SEQ ID NOS: 2699
 SEQ ID NO 1400
 LENGTH: 2744
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-342-887-1400

Alignment Scores:
 Pred. No.: 1,316-150 Length: 2744
 Score: 1359.50 Matches: 301
 Percent Similarity: 40.32% Conservative: 1
 Best Local Similarity: 40.19% Mismatches: 0
 Query Match: 85.02% Indels: 447
 DB: 17 Gaps: 1

US-10-659-004-104 (1-302) x US-10-342-887-1400 (1-2744)

QY 1 MetAlaArgLYLeuSerValIleLeuIleLeuThrPheAlaLeuSerValThrAsnPro 20
 Db 138 ATGGGAGAGAGTATCTGTATCTTGATCTTGACCTTGCCCTCTGTCAACAATCCC 197
 QY 21 LeuHISGLuLeuLYAlaAlaIlePheProGlnThrGluLYLeuSerProAsnTrp 40
 Db 198 CTTCATGAACCTTAAGACAGCTCTTCCCCAGACCACTGAGAAATTAAGTCGAATGG 257
 QY 41 GluSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTyrHisLeuGlnInleu 60
 Db 258 GAATGTGGATTAATGTTGACTTGGCAATTTCCACGCGCATATCATCTCAACAGCTT 317
 QY 61 PheTYrArgTYrGlyGluAsnAsnSerLeuSerValGluGlyPheArgLYLeuLeuGln 80
 Db 318 TTCTACCGCTATGAGAAATAATCTTGTCAATGGAAGGCTTCAGAAATTAATCTTCA 377
 QY 81 AsnIleGlyIleAspLYLeuLYAsnIleLYAsnArgIleHisIleHisAspHisAspHisSer 100
 Db 378 AATATATGGCATATGATTAAGATTAAAGAAATTCATATACCATGATGACACCATCATCTCA 437
 QY 101 AspHisGLuHisHisSerAspHisGLuArgHisSerAspHisGLuHisHisSerAsp--- 119
 Db 438 GACCAAGAGCATCATCATGACATGAGCTCACTCAGACCATGAGCATCATCAGACCAAC 497
 QY 119 ----- 119
 Db 498 GAGCATCACTTGACCATATATCATGCTGCTTGTGTAATAATAGCGAAAGCTTTTGC 557
 QY 119 ----- 119
 Db 558 CCAAGACATGACTCAGATATGTTCAAGTAAGATCTTGAACAACAGCGGGAAGAGCT 617
 QY 119 ----- 119
 Db 618 CACCGACCAAGACATGCCAGTGTAGAAAGAAATGCAAGACAGTGTAGTCTAGTGA 677
 QY 119 ----- 119
 Db 678 GTGACCTCAACTGTGTACAACACTGTCTCTGAAGAACTCACTTCTAGAGACAATAGAG 737

QY 119 ----- 119
 Db 738 ACTCAAGACCTGGAATACTTCCCAAGATGTAAGACAGCTCCATCCACCAAGTCTC 797
 QY 119 ----- 119
 Db 798 ACATCAAGAGCCGGGTGAGCCGGCTGGTGTAGAAAACAATAATCTGTGAGTGAG 857
 QY 119 ----- 119
 Db 858 CCCCCAAAAGCTTTATGATATTCAGAAACAAATGAAAAATCCTCAGAGAGTTTCAT 917
 QY 119 ----- 119
 Db 918 GCATCAAGACTACTGACATCATGCGATGGGCATCCAGGTTCCGGTGAATGCAACAGAG 977
 QY 119 ----- 119
 Db 978 TTCAACTATCTGTGCCAGCCATCATCAACCAATGATGCTAGATCTTGTGTATTCAT 1037
 QY 119 ----- 119
 Db 1038 ACAAGTAAAGAAAGCTGAATCCCTCCAAAGACTATTCATTACAATAGCCTGGGTT 1097
 QY 119 ----- 119
 Db 1098 GGTGTTTATAGCCATTTCATCATGAGTTTCTGTCTGTCTGTGGGGTTATCTTAGTG 1157
 QY 119 ----- 119
 Db 1158 CCTCTCATGATCGGAGTGTTTTCAAATTTCTCTGAGTTTCCTGTGGCACTGGCCGTT 1217
 QY 119 ----- 119
 Db 1218 GGGACTTGAAGTGATGCTTTTTCACCTTTCACATTCATGCAAGTCAACAC 1277
 QY 119 ----- 119
 Db 1278 CATAGCTATGACATGAAGAACGACGAATGGAATGAAGAAAGACCACTTTCAGTCAT 1337
 QY 119 ----- 119
 Db 1338 CTGTCTCTCAAAACATAGAAGAAAGTCTATTTGATTCACAGTGGAGGGTTCACA 1397
 QY 119 ----- 119
 Db 1398 GCTTAGAGGCTGTATTTTCATGTTCTTGTGAAACATGCTTCATCATTTCAACAA 1457
 QY 119 ----- 119
 Db 1458 TTTAAAGATTAAGAAAGAAAGAAATCAGAAAGAACTGAATATGATGATGTGAGATT 1517
 QY 119 ----- 119
 Db 1518 AAGAGCAGTTGTCCAGATATGAATCTCACTTTCACAAATGAGAGAAAGTAGATCA 1577
 QY 119 ----- 119
 Db 1578 GATGATGAACTGAAGGCTATTTACAGACAGATCACAAGACCTCCCATCTTGATTTCT 1637
 QY 119 ----- 119
 Db 1638 CAGCAGCTGAGTCTTGGAAGAAAGAGGCTATGATAGCTCATGCTCATCACAAGAA 1697
 QY 119 ----- 119
 Db 1698 GTCTACATGAATATGTACCCAGAGGTGCAAGATAAATGCCATTCAATTTCCAGCAT 1757
 QY 119 ----- 119
 Db 1758 AACTGGCCAGTACAGAGATCTCATTCACCAACATCATGACTACATCATATTTCCAT 1817
 QY 120 ----- 120
 Db 120 ----- HisHisProHisSerHisSerGlnArgTYrSerArgGluGln 133


```

Db      1818 CATCACCAACCAAAACCAACATCTCTACAGTACACAGCCAGCGTATCTCCGGAGAG 1877
Qy      134 LeuLYsAaPAlaGlyValAlaThrLeuAlaTrpMetValIleMetGlyAspGlyLeuHis 153
Db      1878 CTGAAGAGTGCCTGGCGCTGACCTTTGGCTGATGTGATGAATGAAGGATGAGTGCCTGCAC 1937
Qy      154 AsnHeserAspGlyLeuAlaAlaIleGlyValAlaPheThrGluGlyLeuSerSerGlyLeu 173
Db      1938 AATTTCAGCGAATGCGCTAGCAATGTGTCTCTTTACTGAAGCGTTATCAAGTGTGTA 1997
Qy      174 SerThrSerValAlaValAlaPheCysHisIleGlyLeuProHisGlyLeuGlyAspPheAlaVal 193
Db      1998 AGTACTCTGTGTGCTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2057
Qy      194 LeuLeuLYsAlaGlyMetThrValIleGlyGlnAlaValLeuTYrAsnAlaLeuSerAlaMet 213
Db      2058 CTACTAAAGCGCTGCATGACCGTTAAGCAGCGCTGCTTTATTAATGATGTGTGCAGCCATG 2117
Qy      214 LeuAlaTYrLeuGlyMetAlaThrGlyIlePheIleGlyHisTYrAlaGlyAsnValSer 233
Db      2118 CTGCGGATCTTGGAAATGACACAGAAATTTTCATGTGTCATTAATGCTGAATAATGTTTCT 2177
Qy      234 MetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisValAlaLeuValAspMetVal 253
Db      2178 ATGTGATATTTCGACTTACTGCTGCTGCTTATTCATGATGTGCTGCTGCTGCTGATATGTA 2237
Qy      254 ProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSerArgTrpGlyTYrPhePhe 273
Db      2238 CCTGAATGTGTCACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2297
Qy      274 LeuGlnAsnAlaGlyMetLeuGlyPheGlyIleMetLeuLeuIleSerIlePheGly 293
Db      2298 TTACGAATGCTGGAGATGCTTTGGGTTTGGAAATGATTAATTCATTAATTTGAA 2357
Qy      294 HisLYsIleValAlaPheArgIleAsnPhe 302
Db      2358 CATAAATCTGTCTTCGTAATAATTC 2384

```

RESULT 7
US-10-769-612-1
Sequence 1, Application US/10769612
Publication No.-US20040141983A1

GENERAL INFORMATION:

APPLICANT: Law, Debbie

APPLICANT: Gleib, Kurt C.

APPLICANT: Murray, Richard

APPLICANT: Culp, Patricia

TITLE OF INVENTION: Compositions Against Cancer Antigen L1V-1 And Uses Thereof

FILE REFERENCE: 05882-0129-CPUS08

CURRENT APPLICATION NUMBER: US/10769, 612

PRIOR FILING DATE: 2004-01-29

PRIOR APPLICATION NUMBER: 60/443, 712

PRIOR FILING DATE: 2003-01-29

PRIOR APPLICATION NUMBER: 09/642, 034

PRIOR FILING DATE: 2000-08-18

PRIOR APPLICATION NUMBER: 09/525, 361

PRIOR FILING DATE: 2000-03-15

PRIOR APPLICATION NUMBER: 09/453, 137

PRIOR FILING DATE: 1999-12-02

PRIOR APPLICATION NUMBER: 09/450, 810

PRIOR FILING DATE: 1999-11-29

PRIOR APPLICATION NUMBER: 09/268, 865

PRIOR FILING DATE: 1999-03-15

NUMBER OF SEQ ID NOS: 18

SOFTWARE: Patent version 3.2

SEQ ID NO 1

LENGTH: 2744

TYPE: DNA

ORGANISM: Homo sapiens

US-10-769-612-1

Alignment Scores:

```

Pred. No.: 1,31e-150 Length: 2744
Score: 1359.50 Matches: 301
Percent Similarity: 40.32% Conservative: 1
Best Local Similarity: 40.19% Mismatches: 0
Query Match: 85.02% Indels: 447
DB: 18 Gaps: 1

US-10-659-004-104 (1-302) x US-10-769-612-1 (1-2744)
Qy      1 MetAlaArgLYsLeuSerValIleLeuIleLeuThrPheAlaLeuSerValThrAsnPro 20
Db      138 ATGGCGAGAAATGATCTGTATCTGTATCTGTATCTGTATCTGTATCTGTATCTGTATCT 197
Qy      21 LeuHisGlyLeuLYsAlaAlaPheProGlnThrThrGlyIleSerProAsnTrp 40
Db      198 CTTCATGAACCTAAAGCAGCGCTTTCCCGGACCACTGGAATAATTAATGTCGGAATGG 257
Qy      41 GluSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTYrHisLeuGlnLeu 60
Db      258 GAATCTGGCATTAATGTTGACTTGGCAATTTCCACAGGCAATATCATCTACAAACGCTT 317
Qy      61 PheTYrArgTYrGlyGluAsnAsnSerLeuSerValGlyIlePheArgLYsLeuLeuGln 80
Db      318 TTCTACCGCTATGAGAAAATTAATTTCTTGTCACTGTAAGGGTTCAAGAAATTAATCTTCAA 377
Qy      81 AsnIleGlyIleAspLYsIleLYsArgIleHisIleHisAspHisAspHisSer 100
Db      378 AATTAAGCATATGATTAAGATTAAAGATTCATATACCACTGACACGACCATCATCTCA 437
Qy      101 AspHisGlyHisHisSerAspHisGlyIleArgHisSerAspHisGlyHisSerAsp 119
Db      438 GACCACAGCATCTCACTAGACCATGAGCGTCACTCAAGCATGAGCATCTCAGACCAAC 497
Qy      119 ----- 119
Db      498 GAGCATCACTGTACCATTAATCATGTGCTTGTGTAATAATGAAGAAAGCTTTTGC 557
Qy      119 ----- 119
Db      558 CCAGACATGATCATGATAGTTCAAGTAAGATCTTAAGAAACAGCCAGGGAGAGAGCT 617
Qy      119 ----- 119
Db      618 CACCGACCAACAATGCCAGTGTGAAGAAATGCAAGACAGTGTGTAAGTAA 677
Qy      119 ----- 119
Db      678 GTGACCTCACTGTGTACAACACTGTCTGTGAAGAACTCATTTCTAGACAATAAGAG 737
Qy      119 ----- 119
Db      738 ACTCAAGACCTGGAAAACTTCCCAAGATGTAAGCAGCTCACTCAACCAAGTGTGTC 797
Qy      119 ----- 119
Db      798 ACATCAAAAGACCGGGTGAAGCGGCTGCTGTGAAGAAACAATAATCTGTGAAGAG 857
Qy      119 ----- 119
Db      858 CCGGAAAAAGCTTTATGTAATTCAGAAAACAATAATGAATCTCAAGAGAGTGTTCAT 917
Qy      119 ----- 119
Db      918 GCATCAAAAGCTATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 977
Qy      119 ----- 119
Db      978 TTCACTATCTGTGTCCAGCATCATCAACCAATTAATGATGATGATGATGATGATGATGATG 1037
Qy      119 ----- 119
Db      1038 ACAAGTGAAGAAAGAGGCTGAATAATCCTCAAAAGACTTATTCATAAATAGCTGGGTT 1097

```


QY 41 GluSerGylleapValaspLeuAla11SerThrArg1nTyRhiLeuGlnLeu 60
 Db 258 GAATCTGGCATTTAATGTTGACTTGGCAATTTCCACACGGCAATTCATCTACACAGCTT 317
 QY 61 PheTyArgTyrglygluAsnAsnSerLeuSerVal1gluGlyPheArgLyLeuLeuGln 80
 Db 318 TTTCACCCCTATGAGAAATTAATTTCTTTCAGTTGAGGGTTCAAGAAATTACTTCA 377
 QY 81 Aen11egly1leapLys1lelyArg11eh1s11eh1s1asph1asph1eh1Ser 100
 Db 378 AATTAAGCATATGATTAAGATTAAGATTCATATACCATATACACGACCATCATCTCA 437
 QY 101 Asph1esglu1eh1SerAsph1esgluArg1sSerAsph1esglu1eh1SerAsp--- 119
 Db 438 GACACGAGCATCTCAGACCATGAGCGCTCACTCAGACCATGAGCATCTCAGACAC 497
 QY 119 ----- 119
 Db 498 GAGCATCTGTGACCATATCATGCTGCTTGTGTAAATAATAGCGAAAGCTTTTGC 557
 QY 119 ----- 119
 Db 558 CCAGACCATGACTCAGATAGTTCAAGTAAGTCTTAAACAGCCAGGGAAAGAGACT 617
 QY 119 ----- 119
 Db 618 CACCGACCAAGACATGCCAGTGTAGAGAAATGTCAAGACAGTGTAGTGTAGTA 677
 QY 119 ----- 119
 Db 678 GTGACCTCACTGTGTACACACATGCTCTGTAGAGAACTACTTTCTAGAGCAATAGAG 737
 QY 119 ----- 119
 Db 738 ACTCAAGACCTGGAAAACTTTCCCAAAAGTGTAAAGACAGTCCACTCAACCAAGTGC 797
 QY 119 ----- 119
 Db 798 ACATCAAGAGCCGGGTGAGCCGGCTGGCTGTAGAAAAAAMAATGAATCTGTAGTGA 857
 QY 119 ----- 119
 Db 858 CCCCCAAAAGCTTTATATATTCAGAAAACAAATGAAAAATCCTCAGAGTGTTCAT 917
 QY 119 ----- 119
 Db 918 GCATCAAGCTACTGACATTCATGCGCATGGGCATCCAGTTCCGCTGAAGCAACAGAG 977
 QY 119 ----- 119
 Db 978 TTCAACTATCTCTGTCCAGCATCATCAACAAATGTAGTGTAGATCTTGTCTGATTCAT 1037
 QY 119 ----- 119
 Db 1038 ACAAGTAAAAAGAGCTGAATCCCTCCAAAGACTTATTCATTAACAAATAGCTGGGTT 1097
 QY 119 ----- 119
 Db 1098 GGTGGTTTATAGCATTTCCATCATCATGTTTCCGTCTGTGCGGGGTTATCTTAGTG 1157
 QY 119 ----- 119
 Db 1158 CCTCTCATGAATCGGGTGTCTTTCAAAATTTCTCTGAGTTTCTTTGTGGCACTGGCCGTT 1217
 QY 119 ----- 119
 Db 1218 GGGACTTTGAGTGGTATGCTTTTTCACCTTTCTTCCACATTTCTCATGCAAGTACACAC 1277
 QY 119 ----- 119
 Db 1278 CATAGTCATAGCATGAAGAACAGCAATGGAATGAAAAGAGACACTTTTTCAGTCAT 1337

QY 119 ----- 119
 Db 1338 CTGTCTTCTCAAAACATAGAAAGAGTCCATTATTTTGTATTCACAGTGAAGGCTTAACA 1397
 QY 119 ----- 119
 Db 1398 GCTTAGAGGCGTGTATTTTCATGTTTCTTGTGAACATGTCCTCATGTATGATCAACAA 1457
 QY 119 ----- 119
 Db 1458 TTTAAGATAGAGAAAAAGATACAGAGAACTGAAAAATGATGATGTGGAGATT 1517
 QY 119 ----- 119
 Db 1518 AAGAGCAGTTGCCAAGTATGAATTCACACTTTCAACAAATGAGAGAAAGTATGATCA 1577
 QY 119 ----- 119
 Db 1578 GATGATGAACTGAAGGCTATTTACAGACAGACTCAAGAGACCCCTCCACTTGTATCT 1637
 QY 119 ----- 119
 Db 1638 CAGCAGCCTGAGTCTTGAAGAAAGAGGTATGATAGCTCATGCTCATCCACAGGA 1697
 QY 119 ----- 119
 Db 1698 GTTCAATGAATATGTACCAGAGGTCAGAAATAAATGCCATTCACATTTCCAGAT 1757
 QY 119 ----- 119
 Db 1758 ACACCTGGCCAGTCAGACGATCTCATCCACCATCATGACTACATCATATTCCTCAT 1817
 QY 120 ----- 133
 Db 1818 CATCACCAACCAAAAACACCATCTCTCACAGTACAGCCAGCCCTACTCTCGGAGGAG 1877
 QY 134 Leu1ysAsp1asgluVal1Ala1ThrLeuAla1rPheVal111eMetGlyAspG1yLeu1h1s 153
 Db 1878 CTGAAGATGCGCGCTGCGCATCTGCTGATGTGATGAATAGGATGAGCTTGAC 1937
 QY 154 Asn1SerAspG1yLeuAla11eglyAla1Ala1rPheThrGluGlyLeuSerSerG1yLeu 173
 Db 1938 AATTTCAGCGATGCGCTTACGCAATGGTGTCTCTTACTGAAGGCTTATCAAGTGTTTA 1997
 QY 174 SerThrSerVal1AlaVal1PheCysHisgluLeuProHisgluLeuGlyAspPheAlaVal 193
 Db 1998 AGTACTTCTGTGCTGTGTCTGTATGATGATGCTGCTCATGAATTAGGTGACTTGTCTGTT 2057
 QY 194 Leu1eu1ysAla1gluMetThrVal1lyGlnAla1Val1LeuTyrrAsnAla1euSerAlaMet 213
 Db 2058 CTACTAAAGCTGGCATGACCGTTAAGCAGGCTGTCTTAAATGCAATGTTCAGCCATG 2117
 QY 214 LeuAlaTyrrLeuGlyMetAla1ThrGly11ePhe11eglyHisTyrrAla1gluAsnValSer 233
 Db 2118 CTGGCGTATCTTGGAAATGCAACAGAAATTTTCATTTGCTCTTATGCTGAAAAATCTTCT 2177
 QY 234 MetTyrr1ePheAla1euThrAla1glu1ePheMetHisVal1Ala1euValAspMetVal 253
 Db 2178 ATGTGATATTTGCACTTACTGTGCTTATTCATGATGATGTGCTGTGATGATGATA 2237
 QY 254 ProGluMetLeuHisAsnAspAla1SerAspHisglu1CysSerArgTrrGlyTyrrPhePhe 273
 Db 2238 CCTAAATGCTGCACAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2297
 QY 274 LeuGlnAsnAla1gluMetLeuGly1PheGly11eMetLeuLeu11eSer11ePheGlu 293
 Db 2298 TTACAGAAATGCTGGATCTTTTGGGTTTGGATTAATGTTACTTATTTCCATTTTGA 2357
 QY 294 HisLy11eVal1PheArg11eAsn1Phe 302
 Db 2358 CATAAATCGTGTTCGTATAAATTTTC 2384

RESULT 9

```
US-10-198-846-10007
; Sequence 10007, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10007
; LENGTH: 3537
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 2, 3533, 3534, 3535, 3536, 3537
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-10007

Alignment Scores:
Pred. No.: 1,92e-150 Length: 3537
Score: 1359.50 Matches: 301
Percent Similarity: 40.32% Conservative: 1
Best Local Similarity: 40.19% Mismatches: 0
Query Match: 85.02% Indels: 447
DB: 14 Gaps: 1

US-10-659-004-104 (1-302) x US-10-198-846-10007 (1-3537)
QY 1 MetAlaAGlyLeuSerValIleLeuTIRPheAlaLeuSerValThrAspPro 20
DB 138 ATGGGAGAGATTTCGTATCTTGATCCGACCTTCCCTCTCTGTCACAATCCC 197
QY 21 LeuHISGluLeuAlaAlaIlePheProGlnThrGluLeuIleSerProAsnTrp 40
DB 198 CTTCACTGAACCTAAAGCAGCTGCTTCCCGACCACTGAGAAATTAAGTCGATTGG 257
QY 41 GluSerGlyIleAsnValAspLeuAlaIleSerThrArgIleGluGlnGlnLeu 60
DB 258 GAATCTGGCATTAATGTGACTTGGCAATTTCCACAGCGCAATATCATCTCAACAGCTT 317
QY 61 PheTyrArgTyrGlyGluAsnAsnSerLeuSerValIleGluIlePheArgLysLeuGln 80
DB 318 TTCACCCCTAAGGAAATTAATCTTTGTCAAGTTGAAGGTTAGAAAATTAATCTTCAA 377
QY 81 AsnIleGlyIleAspLysIleLysArgIleHisIleHisIleAspHisAspHisSer 100
DB 378 AATATAGGCATGATTAAGATTAAAGATCCATATACACATGACCAAGACCATCATCTCA 437
QY 101 AspHisGluHisHisSerAspHisGluArgHisSerAspHisGluHisHisSerAsp 119
DB 438 GACCAAGAGCATCACTCAGACCATGAGCGTCACTCAGACCATGACATCATCAGACAC 497
QY 119 ----- 119
DB 498 GAGCATCACTGACCATTAATCATGCTGCTGTGTAATAATAGCGAAAGCTTTTGC 557
QY 119 ----- 119
DB 558 CCAAGCATGACTCAGATTAGTTCAAGTAAAGATCTTAGAAACAGCCAGGGAAAGAGACT 617
QY 119 ----- 119
DB 618 CACGACCAAGAACATGCCAGTGTAGAAAGAAATGTCAAGACAGTGTAGTGTAGTAA 677
```

```
QY 119 ----- 119
DB 678 GTGACCTCAACTGTGTACAACTGTCTGAAAGAACTCATCTTTAGAGACAATAGAG 737
QY 119 ----- 119
DB 738 ACTCAAGACCTGGAATACTTCCCAAGATGTAGACAGCTCCACTCCACCAAGTGC 797
QY 119 ----- 119
DB 798 ACATCAAGAGCCGGGTGAGCCGGCTGCTGTAGAAAACAAATGAATCTGTAGTGTAG 857
QY 119 ----- 119
DB 858 CCCCAGAAAGCTTTATGTATTCAGAAACAAATGAAATTCCTCAGAGAGTGTTCAT 917
QY 119 ----- 119
DB 918 GCATCAAGCTACTGACATCTCATGCGATGGCATCCAGGTTCCGCTGAATGCAACAGAG 977
QY 119 ----- 119
DB 978 TTCAACTATCTGTGTCCAGCATCATCAACCAATGTAGCTAGATCTGTGTATTATCAT 1037
QY 119 ----- 119
DB 1038 ACAAGTGAAGAGGCTGAATCCCTCCAAAGCTATTATCAAAATAGCCTGGGTT 1097
QY 119 ----- 119
DB 1098 GGTGTATTATAGCATTTTCATCATCATGATTCCTGTCTGTGGGGTTATCTTAGTG 1157
QY 119 ----- 119
DB 1158 CTTCTCATGAATCGGAGTGTTCCTCAAAATTCCTCTGAGTTCTCTGTGGCACTGGCGGT 1217
QY 119 ----- 119
DB 1218 GGGACTTGAGTGTATGCTTTTTCACCTTCTTCCACATTTCTCATGCAAGTCACAC 1277
QY 119 ----- 119
DB 1278 CATAGTATAGCATGAAGAACCAAGATGAATGAAGAGAGCACTTTGATGTCAT 1337
QY 119 ----- 119
DB 1338 CTGTCTTCTCAAAACATAGAAAGAGTCTATTGTTGATTCACGTGAAAGGTCTAACA 1397
QY 119 ----- 119
DB 1398 GCTTAGAGGCTGTATTTTCATGTTTCTTGTGTAACATGCTTCACATTTGATCAACAA 1457
QY 119 ----- 119
DB 1458 TTTAAGATGAAGAAAGAAAGATCAGAGAAACCTGAAATATGATGATGTGAGATT 1517
QY 119 ----- 119
DB 1518 AAGAGAGTGTGCCAAGTATGATTCCACTTTCAACAAATGAGAGAAAGTAGATACA 1577
QY 119 ----- 119
DB 1578 GATGATGAAGAGCTATTATTCGAGCAGACTCACAAGAGCCCTCCACTTTGATTCT 1637
QY 119 ----- 119
DB 1638 CAGAGCGTGAAGTCTTGAAGAAAGAGGTATGATGATGATGATGATGATGATGATGAT 1697
QY 119 ----- 119
DB 1698 GTTCAATGATATATGTATCCAGAGGTGCAAGAAATTAATGCAATTGATTCAGAGAT 1757
```

```

Oy 119 ----- 119
Db 1758 AACACTGGCAGTACGAGCATCTCATTCACACCATCATGACTACATCATATTTCAT 1817
Oy 120 ----- HiHiPProHiseRHisSerGlnArgTyrSerArgGlu 133
Db 1818 CATCACCAACCAAAACACCATCTCTCAAGTCCACACCGCTACTCTCGGAGGAG 1877
Oy 134 LeuYsaAaPaJaJaValAlaThrLeuAlaTPrMetValIleMetGlyAspGlyLeuHis 153
Db 1878 CTGAAGAGTCCGGGGTGGCCACTGGCTGGATGGATGGATGGATGGATGGATGGATGG 1937
Oy 154 AenPseSerAaPaJaJaValAlaThrLeuAlaTPrMetValIleMetGlyAspGlyLeu 173
Db 1938 AATTTCACGATGGCTGACATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 1997
Oy 174 SerThrSerValAlaValAlaPheCysHisGluLeuProHisGluLeuGlyAspPheAlaVal 193
Db 1998 AGTACTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2057
Oy 194 LeuLeuYsaAaPaJaJaValAlaThrLeuAlaTPrMetValIleMetGlyAspGlyLeu 213
Db 2058 CTACTAAAGGCTGGGATACCGTTAAGCAGCTGCTCTTAAATGATGATGATGATGATG 2117
Oy 214 LeuAlaTyrLeuGlyMetAlaThrGlyIlePheIleGlyHisTyrAlaGlyAsnValSer 233
Db 2118 CTGGCGTATCTTGGAAATGGCAACAGAAATTTTCATGATTCATGATTCATGATTCAT 2177
Oy 234 MetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisValAlaLeuValAspMetVal 253
Db 2178 ATGGAGATATTTGCACTTACTGCTGCTTATTCATGATTCATGATTCATGATTCATG 2237
Oy 254 ProGluMetLeuHisAenAaPaJaJaSerAspHisGlyCysSerArgTyrGlyTyrPhePhe 273
Db 2238 CCTGAAGGCTGCACAAATGATGATGATGATGATGATGATGATGATGATGATGATG 2297
Oy 274 LeuGlnAenAlaGlyMetLeuLeuGlyPheGlyIleMetLeuLeuIleSerIlePheGlu 293
Db 2298 TTACGAATGCTGGGATCTTTGGGTTTGGAAATTAATTAATTAATTAATTAATTA 2357
Oy 294 HisLysIleValPheArgIleAsnPhe 302
Db 2358 CATAAATCGTGTTCGTATAAATTTC 2384

RESULT 10
US-10-188-832-176
; Sequence 176, Application US/10188832
; Publication No: US20040076955A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Aziz, Natabha
; TITLE OF INVENTION: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder
; FILE REFERENCE: 018501-0023330US
; CURRENT APPLICATION NUMBER: US/10/188, 832
; PRIOR FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/302,814
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/310,099
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/343,705
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 176
; LENGTH: 2268
; TYPE: DNA

```

```

; ORGANISM: Homo sapiens
US-10-188-832-176
Alignment Scores:
Pred. No.: 2,236-150 Length: 2268
Score: 1356.50 Matches: 301
Percent Similarity: 40.00% Conservative: 1
Best Local Similarity: 39.87% Mismatches: 0
Query Match: 84.83% Indels: 453
DB: 17 Gaps: 1

US-10-659-004-104 (1-302) x US-10-188-832-176 (1-2268)
Oy 1 MetAlaArgLysLeuSerValIleLeuIleLeuThrPheAlaLeuSerValThrAspPro 20
Db 1 ATGGGAGAGAAATTAATCTGTAATCTGTAATCTGTAATCTGTAATCTGTAATCTGTA 60
Oy 21 LeuHisGluLeuYsaAaPaJaJaValAlaPheProGlnThrThrGlyLysIleSerProAsnTrp 40
Db 61 CTTCATGACTTAAAGCAGCTGCTTCCCGACACCACTGAGAAATTAATGATCCGATGG 120
Oy 41 GluSerGlyIleAenValAspLeuAlaIleSerThrArgGlnTyrHisLeuGlnGlnLeu 60
Db 121 GAATCTGGCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
Oy 61 PheTyrArgTyrGlyGluAenAaPaJaJaSerLeuSerValGluGlyPheArgLysLeuGln 80
Db 181 TTCTACCGCTAATGAGAAATTAATTTCTTGTCAATGAGGTTTCAGAAATTAATTTCT 240
Oy 81 AsnIleGlyIleAaPlyHisLeuArgIleHisIleHisHisAspHisAspHisHisSer 100
Db 241 AATTAATGAGCATGATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 300
Oy 101 AspHisGluHisHisSerAspHisGluArgHisSerAspHisGluHisHisSerAsp 119
Db 301 GACCACGAGCATCACTGACACCATGAGGCTCAGCCATGAGCATCACTCAGACCAAC 360
Oy 119 ----- 119
Db 361 GAGCATCACTGACCATGATCATCACTCCACCATATATATGCTTCTGTGTAATAAT 420
Oy 119 ----- 119
Db 421 AAGCGAAAGCTCTTTGCCAGACCATGATCACTCAATGATGATGATGATGATGATG 480
Oy 119 ----- 119
Db 481 AGCAGGGGAAAGAGGCTCAACCGACCAAGAACATGCAATGATGATGATGATGATG 540
Oy 119 ----- 119
Db 541 AGTGTAGTGTAGTGAAGTGAAGTCACTCACTGTGTACAACTGTCTGTGAAGAACTC 600
Oy 119 ----- 119
Db 601 TTTCTAGACATATAGAGACTCAAGACTGGAAAACTTTCCCAAAATGTATGACAGC 660
Oy 119 ----- 119
Db 661 TCCACTCAACCACTGTCATCACTAAAGAGCCGGGTGAGCGGCTGGCTGTAGGAAAA 720
Oy 119 ----- 119
Db 721 AATGAATCTGTAGTGAAGCCCGAAAAAGCTTTATGTATTCAGAAACAAATGAAAA 780
Oy 119 ----- 119
Db 781 CTCGAGAGTGTTCATATGATCAAGACTACTGATCTATGATGATGATGATGATGAT 840
Oy 119 ----- 119
Db 841 CCGCTGAATGCAACAGAGTTCACTATCTGTCTGACAGCATCATCAACAAATGATGCT 900

```

```

QY 119 ----- 119
Db 901 AGATCTGTCTGATTCATACAAAGTGAAGAGGCTGAAATCCCTCCAAAGACCTATTCA 960
QY 119 ----- 119
Db 961 TTACAAATAGCGTGGGTTGGTGGTTTATAGCCATTTCATCATCATGTTTCCTGTCTGTG 1020
QY 119 ----- 119
Db 1021 CTGGGGTTATCTTAGTGCCTCTCATGAATCGGGTGTGTTTCAATTTCTCTGAGTTTC 1080
QY 119 ----- 119
Db 1081 CTTGTGGACCTGGCCGTTGGACCTTTGAGTGTGATGCTTTTTCACCTCTTCCACAT 1140
QY 119 ----- 119
Db 1141 TCTCATGCAAGTCACACCATATGATAGCCATGAAGAACCAGCAATGGAAATGAAGAA 1200
QY 119 ----- 119
Db 1201 GGACACATTTTCATCTCTGTCTCTCAAAAATAGAAAGTGCCTATTGTTGATTC 1260
QY 119 ----- 119
Db 1261 ACGTGAAGGCTCTAACAGCTCTAGAGGCTGTATTTTCATGTTCTTGTGACATGTC 1320
QY 119 ----- 119
Db 1321 CTCACATTGATCAAAACAATTAAAGTAAAGAAAGAAAGAAATCAGAAACCTGAAAT 1380
QY 119 ----- 119
Db 1381 GATGATGATGAGATTAAAGAGAGTTTCCAATGATGATCTCAATTTCACAAAT 1440
QY 119 ----- 119
Db 1441 GAGGAGAAAGTATACAGATATGAACTGAAGGCTATTTCAGAGAGACTCACAAGAG 1500
QY 119 ----- 119
Db 1501 CCTCCACTTGAATTCAGAGAGCTGCAAGTCTTGAAAGAAAGAGGTATGATAGCT 1560
QY 119 ----- 119
Db 1561 CATGCTCATCAAGAGAGCTTACATGATATGTACCCAGAGGTGCAAGATTAATGC 1620
QY 119 ----- 119
Db 1621 CATTCACATTTCCAGATACACTCGGCCAGTCAGAGCATCTCATTCACCATCATGAC 1680
QY 120 ----- 127
Db 1681 TACCATCATATTTCTCATCATCAACACCAAAACCAATCTTCACAGTCAAGCCAG 1740
QY 128 ArgTyrSerArgGluGluLeuLysAspAlaGlyValAlaThrLeuAlaTrpMetValIle 147
Db 1741 CCTTACTCTCGGAGAGAGCTGAAGATGCGGCGTCCGACATTTGGCCCTGAGTGGGAA 1800
QY 148 MetGlyAspGlyLeuHisAsnProSerAspGlyLeuAlaIleGlyValAlaThrArgIu 167
Db 1801 ATGGGTGATGGCTGCAAAATTCACAGCATGGCTGACAAATGGTGTCTCTTTACTGA 1860
QY 168 GlyLeuSerSerGlyLeuSerThrSerValAlaValPheCysHisGlyLeuProHisGlu 187
Db 1861 GCGTATATCAAGTGGTTTAGTACTTCTGTGCTGTGTCTGTGATAGTATGGCTATGA 1920
QY 188 LeuGlyAspPheAlaValLeuLeuLysAlaGlyMetThrValIleGlnAlaValLeuTyr 207
Db 1921 TTAGGTGATCTTGTCTTCTACTAAAGCTGGGATGACCGTTTAAAGCAGGCTGTCTTAT 1980
QY 208 AsnAlaLeuSerAlaMetLeuAlaTyrLeuGlyMetAlaThrGlyLeuPheIleGlyHis 227

```

```

Db 1981 AATGATTCAGAGCATGCTGGGTATCTTGAATGCAACAGGAATTTTCATGTCAT 2040
QY 228 TyrAlaGluAsnValSerMetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisVal 247
Db 2041 TATGCTAAATGATGTTTATGATGATATTTGACCTTACTCTGCTTATTCATATGATGT 2100
QY 248 AlaLeuValAspMetValProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSer 267
Db 2101 GCTTGGTGTATATGATGATCTGAAATGCTGCAATATGCTAGTACATGATGATGATG 2160
QY 268 ArgTrpGlyTyrPhePheLeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeu 287
Db 2161 CGCTGGGGATATTTCTTTTAAAGAAATGCTGGATGCTTTTGGGTTTGGAAATATGTTA 2220
QY 288 LeuIleSerIlePheGluHisLysIleValPheArgIleAsnPro 302
Db 2221 CTATATTCATATTTGAACATAAATGCTGTGTTGTTAAATTTTC 2265

```

RESULT 11

```

US-10-659-004-105
; Sequence 105, Application US/10659004
; Publication No. US20050048507A1
; GENERAL INFORMATION:
; APPLICANT: Zhong et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-608
; CURRENT FILING DATE: 2003-09-09
; PRIOR APPLICATION NUMBER: US/10/659, 004
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/295,607
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/295,661
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/296,404
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/296,418
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/297,414
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 60/297,567
; PRIOR FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 60/298,285
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/298,556
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/299,949
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: 60/300,883
; PRIOR FILING DATE: 2001-06-26
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: Curaseqdist version 0.1
; SEQ ID NO 105
; LENGTH: 2268
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2265)
US-10-659-004-105

```

```

Alignment Scores:
Pred. No.: 2,23e-150 Length: 2268
Score: 1356.50 Matches: 301
Percent Similarity: 40.00% Conservative: 1
Best Local Similarity: 39.87% Mismatches: 0
Query Match: 84.83% Indels: 453
De: 19 Gaps: 1

```

US-10-659-004-104 (1-302) x US-10-659-004-105 (1-2268)

QY 1 MetAlaArgLysLeuSerValIleLeuIleLeuThrPheAlaLeuSerValIleThrAsnPro 20

Db	1	ATGGGAGGAAAGTATCTGTATCTTGATCTTGATCTTGACCTTTGGCCCTCTCTGTACAAACCC	60
Qy	21	LeuH1eGIuLeuLyA1a1a1a1aPheProG1nThrG1uLyA1eSerProaSerTrp	40
Db	61	CTTCATGAACCTPAAAAAGCAGCTGCTTTCCCCAGACACCTGAGAAAATTATGTCGAATTGG	120
Qy	41	GIuSerGIY11eAenValAaPheLeuA1a1eSerThraArg1nThrHiLeuGIu1nLeu	60
Db	121	GAATCTGGCATTTAAGTTGACTTGCGCAATTTCCACAGCGCAATATCATCTCAACAGGCTT	180
Qy	61	PheTyraArgTyrg1yGIuAaPnaSerLeuSerValGIuGIyPheArgLyLeuLeuGIu	80
Db	181	TTCTACCCCTATGAGAAAATAATTCTTTGTCACTTGAGGGGTTCAAGAAAATTACTTCA	240
Qy	81	Aen11eGIY11eAerLyse11eLyAArg11eHiS11eHiAaPna1aAerHi1aSer	100
Db	241	AATATAGGCATATGATATTAATTTAAAGAAATCCATATACACACAGACCATCATCTCA	300
Qy	101	AerHi1aGIuHiA1eSerAaPna1eGIuAArgHi1eSerAaPna1eGIuHi1aSerAaP	119
Db	301	GACCAAGACATCACTCAGACCATGAGCGTCACTCAGACCATGACATCATCAGACGAC	360
Qy	119	-----	119
Db	361	GAGCATCACTGACCATGATCATCATCTCAACATATCATGCTTCTGTAAAAAT	420
Qy	119	-----	119
Db	421	AAGCGAAAAGCTTTTGGCCAGACCATGATCATGATTAAGTTAGTAAGATCTTGAAG	480
Qy	119	-----	119
Db	481	AGCCAGGGGAAAAGAGCTCACCGACAGAAATGCCAGTGTAGAGAAATGTCAAGAC	540
Qy	119	-----	119
Db	541	AGTGTAGTGTAGTAAGTGAAGTGAAGTCACTCACTGTGTACAAACATGTCTTGAAAGAACTCAC	600
Qy	119	-----	119
Db	601	TTTCTAGAGACAAATAGAGACTCCAGACCTGGAAAATCTTCCCAAGATGTAGACAGC	660
Qy	119	-----	119
Db	661	TCCACTCCACCCAGTGTACATCAAAAGCCGGGTGAGCCGGCTGGTGTAGGAAAA	720
Qy	119	-----	119
Db	721	AATGAATCTGTAGTGAAGCCCGAAAAGGCTTTATGTATTCAGAAAACAAATGAAT	780
Qy	119	-----	119
Db	781	CCTGAGAGTGTTCATATGATCAAAAGCTACTGATCATGTGATGGATGGGATCCAGGTT	840
Qy	119	-----	119
Db	841	CCGCTGAATGCACAGAGTTCAACTATCTGTGTCCAGGCATCATCAACCAATGATGCT	900
Qy	119	-----	119
Db	901	AGATCTGTGTATTCATACAAAGTAAAAAGAGCTGAATTCCTCCAAAGACTTATTC	960
Qy	119	-----	119
Db	961	TTACAATAGCCTGGGTGTGTGTATTATAGCCATTTCCATCATCATGTTTCTGTCTGTG	1020
Qy	119	-----	119
Db	1021	CTGGGGTTATCTTAGTGCTCTCATGAATCGGTTGTTTTCAAATTTCTCTGAGTTTC	1080
Qy	119	-----	119
Db	1081	CTTGTGGACCTGGCGGTGGGACTTGTAGTGTGATGCTTTTTCACACTTCTTCCACAT	1140
Qy	119	-----	119
Db	1141	TCTCATGCAAGTCAACACCATATGATCATATGACATGAAAGACGACAAATGAAAA	1200
Qy	119	-----	119
Db	1201	GGACCATTTTATGATCATCTGTCTTTCAAAACATAGAAAGAGTGCCTATTTTGATTC	1260
Qy	119	-----	119
Db	1261	ACGTGAAGGGTCTTAACAGCTCTAGAGGCTGATTTTCATGTTTCTTGTGAACATGTC	1320
Qy	119	-----	119
Db	1321	CTCACATGATCAACAATTTAAAGATAGAAAGAAAAGAAATCAGAGAAAACCTGAAAT	1380
Qy	119	-----	119
Db	1381	GATGATGATGTGAGATTAAAGAGCATTTGTCCAAAGTATGAAATCTCAACTTCAACAAAT	1440
Qy	119	-----	119
Db	1441	GAGGAGAAAGTATACAGATGATCGAACTGAAGGCTATTTACAGACAGACTCACAGAG	1500
Qy	119	-----	119
Db	1501	CCCTCCACTTTGATTTCTCAGACAGCTGCACTTTGGAAGAAAGAGTCAATAGTACT	1560
Qy	119	-----	119
Db	1561	CATGCTATCCACAGAAAGTCTCAATGATATGTATCCACAGAGGTGCAAGATTAATGCC	1620
Qy	119	-----	119
Db	1621	CATTACATTTCCAGATACACTGGGCCAGTCAGACATCATTTACACCATCATGAC	1680
Qy	120	-----His1aProH1eSerH1eSerG1n	127
Db	1681	TACCATCATATTTCTCCATCATCACCAACCAACCACTCCACAGTCCACAGTCCAG	1740
Qy	128	ArgTyraArgGIuGIuLeuLyAaPna1aGIY1a1aThrLeuA1aThrMetVal11e	147
Db	1741	CGTACTCTCGGAGAGCTGAAAGATGCCGCGCTGCACTGTGCTGATGATGTGTAT	1800
Qy	148	MetGIYAerGIYLeuHiAaPnaSerAaPna1eGIY1a1aGIY1a1aPheThrGIu	167
Db	1801	ATGGGTGATGCTGACCAATTTCAAGCATGGCTTAGCAATTTGGTGTCTTTTACTGA	1860
Qy	168	GIYLeuSerSerGIYLeuSerThrSerVal1aValPheCyHiGIuLeuProH1eGIu	187
Db	1861	GGCTTATCAAGTGTATTAAGTACTTGTGTCTGTGTCTGTCAATGAGTCCCTCATGA	1920
Qy	188	LeuGIYAaPnaPheA1aValLeuLyA1aGIY1a1aMetThrValYsGIuA1aValLeuYr	207
Db	1921	TTAGGTGACTTGTCTGCTTACTAAAGCTGACATGACCGTTAAGCAGCGTGTCTTAT	1980
Qy	208	AaA1aLeuSerA1aMetLeuA1aTyrlLeuGIY1a1aMetA1aThrGIY11ePheH1eGIYHiS	227
Db	1981	AATGATTTGTACGACCATCTGCGATCTTGTGAATGCAACAGAAATTTTCATGCTCAT	2040
Qy	228	TyrA1aGIuAaPna1aSerMetTrp11ePheA1aLeuThra1aGIYLeuPheMetHiSVal	247
Db	2041	TATGTGAAAATGTTTCTATGTGATATTTCCATCTGCTGCTTATATCATGATATGTT	2100
Qy	248	AlaLeuValAaPnaMetValProGIuMetLeuHiAaPnaPna1aSerAaPna1eGIYCySer	267
Db	2101	GCTCTGTTGATATGATGATCTGAAATGCTGCAATGATGCTAGTACATGATGATGATG	2160
Qy	268	ArgTrpGIYTyrlPhePheLeuGIuAaPna1aGIY1a1aMetLeuGIY1a1aMetLeu	287
Db	2161	CGCTGGGGTATTTCTTTTACAGAAATGCTGGATGCTTTTGGGTTTGGAAATATATGTTA	2220

Qy 288 LeuileserilepneGluHieLysIleValpheArgIleAsnPro 302
Db 2221 CTTATTTCATATTGAGACATAAATCGTGTTCGTATAAATTC 2265

RESULT 12

US-10-295-027-55
Sequence 55, Application US/10295027
Publication No. US2003023250A1
GENERAL INFORMATION:

APPLICANT: Afar, Daniel
APPLICANT: Aziz, Natasha
APPLICANT: Ginsberg, Wendy M.
APPLICANT: Gish, Kurt C.
APPLICANT: Glynn, Richard
APPLICANT: Hevezi, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard
APPLICANT: Watson, Susan R.
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of diagnosis of Cancer, Compositions and
TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
FILE REFERENCE: 018501-012500US
CURRENT APPLICATION NUMBER: US/10/295,027
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/663,733
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,394
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/332,464
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/334,393
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/340,376
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/347,211
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US 60/347,349
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/355,250
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/356,714
PRIOR FILING DATE: 2002-02-13
Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 1386

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 55

LENGTH: 3461

TYPE: DNA

ORGANISM: Homo sapiens

US-10-295-027-55

Alignment Scores:

pred. No.: 4 24e-150 Length: 3461
Score: 1356.50 Matches: 301
Percent Similarity: 40.00% Conservative: 1
Best Local Similarity: 39.87% Mismatches: 0
Query Match: 84.83% Indels: 453
DB: 17 Gaps: 1

US-10-659-004-104 (1-302) x US-10-295-027-55 (1-3461)

Qy 1 MetaIaArgIysLeuSerValIleLeuIleLeuThrPheAlaLeuSerValThrAsnPro 20
Db 138 ATGGCGAGGAAGTATCTGTATCTGTATCTGATCTGACCTTCTCTCTGTACAAATCCC 197
Qy 21 LeuHieGluLeuLysAlaIlePheProGlnThrGluLysIleSerProAsnTrp 40
Db 198 CTTTCATGAACTAAAGAGCGCTTCTCCCGAGCCCTAGAAATTAAGTCGAATTGG 257
Qy 41 GluSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTyrHisLeuGlnGlnLeu 60

Db 258 GAATCTGCATTAATGTGATCTTGCAATTTCCACAGCGCAATATCATCAACAGCTT 317
Qy 61 PheTYArgTYrGlyGluAsnAsnSerLeuSerValGluGlyPheArgLysLeuLgln 80
Db 318 TTCTACCGCTATGAGAAATAATTTCTTGTCTGAGTGAAGGTTCCAGAAATTAATCTCAA 377
Qy 81 AsnIleGlyIleAspLysIleLysArgIleHisIleHisIleAspHisAspHisSer 100
Db 378 AATATAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 437
Qy 101 AspHisGluHisIleSerAspHisGluArgHisSerAspHisGluHisSerAsp 119
Db 438 GACACAGCATCTACAGACCATGAGCGTCACTCAGACCATGATGATCACTCAGACAC 497
Qy 119 ----- 119
Db 498 GAGCATCTGTGACCATGATCATCACTCCACCATATATATGCTGCTTGTGTAATAAT 557
Qy 119 ----- 119
Db 558 AAGGAAAGCTCTTGTCCAGACCATGATGATGATGATGATGATGATGATGATGATGATGAT 617
Qy 119 ----- 119
Db 618 AGCAGGGAAGAGAGCTCAACGACAGAAATGCGAGTGAAGAAATGTCAGAGAC 677
Qy 119 ----- 119
Db 678 AGTGTAGTGTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 737
Qy 119 ----- 119
Db 738 TTTCTAGAGCAATAGAGACTCAAGACTGGAAGAACTTTCCCAAGATGAAGAGC 797
Qy 119 ----- 119
Db 798 TCCATCCACCAAGTGTACATCAAGAGCGGAGCGGCTGCTGTAGAGAAACA 857
Qy 119 ----- 119
Db 858 AATGATGTGTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 917
Qy 119 ----- 119
Db 918 CCTGAGAGTGTTCATGATCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 977
Qy 119 ----- 119
Db 978 CCGCTGAATGCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1037
Qy 119 ----- 119
Db 1038 AGATCTGTGTATTCATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1097
Qy 119 ----- 119
Db 1098 TTACAATAGCTGAGTGTGTGTATTAAGCATTTTCATCATCATCATCATCATCATCATCATCAT 1157
Qy 119 ----- 119
Db 1158 CTGGGGGTATTAATGAGT 1217
Qy 119 ----- 119
Db 1218 CTTGTGAGCATGCGCGTGTGAGCTTTTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1277
Qy 119 ----- 119
Db 1278 TCTCATGAAAGTCAACACCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1337
Qy 119 ----- 119
Db 1338 GAGCACTTTTGAATCTGTCTTCAAAATAGAAAGAGTCTATTTTGAATTC 1397

Qy	119	-----	119
Db	1398	ACGTGGAAAGGCTTAACAGCTCTAGAGGCGCTGATTTTCATGTTTCTTGTTGAACATGTC	145
Qy	119	-----	119
Db	1458	CTCACATTGATCAACAATTTTAAGTAAGTAAGAAAGAAAGAAAGTAATCGAAGAAAGAACTGTGAAT	151
Qy	119	-----	119
Db	1518	GATGATGATGTGGAGATTAAAGAGAGTTGTCCAGTATGAATTCACATTTCAACAAAT	157
Qy	119	-----	119
Db	1578	GAGGAGAAAGTAGATACAGATGATGAACTGAAGGCTATTACAGACAGACTCACAGAG	163
Qy	119	-----	119
Db	1638	CCCTCCCACTTTGATTTCTCAGCAGCGCTGACGTTTGGAAGAAAGAGGCTATGATAGCT	169
Qy	119	-----	119
Db	1698	CATGCTCATCCACAGGAAGTCTACATGAATATGTACCAGAGGAGTGCAAGATAATGC	175
Qy	119	-----	119
Db	1758	CATTCACTTTCCACAGATACACTCGGCGCAGTACAGACGATCTCATTCACCACTCATGAC	181
Qy	120	-----HisH1sProHisSerHisSerGln	127
Db	1818	TACCATCATATTTCTCCATCATGACCCACCAACAAACACCATCTCTCAGTCAGTCAAGCAG	187
Qy	128	ArgTYrSerAArgGluGluLeuLysAspAlaGlyValAlaThrLeuAlaIrrPheValIle	147
Db	1878	CGCTACTCTCGGAGAGAGCTGMAAGATGCCGCGCTGCACACTTGGCCTGATGGTGATA	193
Qy	148	MetGlyAspGlyLeuHisAsnPheSerAspGlyLeuAlaIleGlyValAlaIrrPheTrgIu	167
Db	1938	ATGGGTGATGGCCTCAGACATTTCCAGCGATGGCCTTACGAAATGGTCTCTTTACTGAA	199
Qy	168	GlyLeuSerSerGlyLeuSerThrSerValAlaValPheCysHisGluLeuProHisGlu	187
Db	1998	GGCTTATCAAGTGGTTTAAGTACTTCTGTGCTGCTGTCTGTATCATAGATGGCTCATGAA	205
Qy	188	LeuGlyAspPheAlaValLeuLeuLysAlaGlyMetThrValLysGlnAlaValLeuTy	207
Db	2058	TTAGGTGACTTGTGCTCTACTTAAAGGCTGGCATGACCGTTAAGCAGGCTGCTCTTAT	211
Qy	208	AsnAlaLeuSerAlaMetLeuAlaTyrrLeuGlyMetAlaThrGlyIlePheIleGlyHis	227
Db	2118	AATGATTTGTCAAGCAGTCTGGCGATCTTTGGAAATGGCAACAGAAATTTTCATTGGTCA	217
Qy	228	TyrAlaGluAsnValSerMetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisVal	247
Db	2178	TATGCTGAAAAATGTTTCTATGTGATATTTGCACTTACTGCTGCTTATTCATGATGTT	223
Qy	248	AlaLeuValAspMetValProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSer	267
Db	2238	GCTCGTGGTGAATGTGATCTGTAATAATGTCGCACAAGATGCTAGTACCATGATGTAC	229
Qy	268	ArgTYrGlyTyrrPhePheLeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeu	287
Db	2258	CGCTGGGGGATATTTCTTTTACGAATGCTGGGATCTTTTGGGTTTGGAAATATGTTA	235
Qy	288	LeuIleSerIlePheGluHisIleLysIleValPheArgIleAsnPhe	302
Db	2358	CTTATTTCCATATTTGAACATATAAATCGTGTTTCGATATAAATTC	2402

```

: GENERAL INFORMATION:
: APPLICANT: Afar, Daniel
: APPLICANT: Aziz, Natasha
: APPLICANT: Gineberg, Wendy M.
: APPLICANT: Gish, Kurt C.
: APPLICANT: Glyne, Richard
: APPLICANT: Hevezi, Peter A.
: APPLICANT: Mack, David H.
: APPLICANT: Murray, Richard
: APPLICANT: Watson, Susan R.
: APPLICANT: Eos Biotechnology, Inc.
: TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
: FILE REFERENCE: 018501-012500US
: CURRENT FILING DATE: 2002-11-13
: PRIOR APPLICATION NUMBER: US 09/663,733
: PRIOR FILING DATE: 2000-09-15
: PRIOR APPLICATION NUMBER: US 60/350,666
: PRIOR FILING DATE: 2001-11-13
: PRIOR APPLICATION NUMBER: US 60/335,394
: PRIOR FILING DATE: 2001-11-15
: PRIOR APPLICATION NUMBER: US 60/332,464
: PRIOR FILING DATE: 2001-11-21
: PRIOR APPLICATION NUMBER: US 60/334,393
: PRIOR FILING DATE: 2001-11-29
: PRIOR APPLICATION NUMBER: US 60/340,376
: PRIOR FILING DATE: 2001-12-14
: PRIOR APPLICATION NUMBER: US 60/347,211
: PRIOR FILING DATE: 2002-01-08
: PRIOR APPLICATION NUMBER: US 60/347,349
: PRIOR FILING DATE: 2002-01-10
: PRIOR APPLICATION NUMBER: US 60/355,250
: PRIOR FILING DATE: 2002-02-08
: PRIOR APPLICATION NUMBER: US 60/356,714
: PRIOR FILING DATE: 2002-02-13
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 1386
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 838
: LENGTH: 3461
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-295-027-838
Alignment Scores:
Pred. No.: 4,24e-150 Length: 3461
Score: 1356.50 Matches: 301
Percent Similarity: 40.00% Conservative: 1
Best Local Similarity: 39.87% Mismatches: 0
Query Match: 84.83% Indels: 453
DB: 17 Gaps: 1
US-10-659-004-104 (1-302) x US-10-295-027-838 (1-3461)
QY 1 MetaIaArgLYsLeuSerValIleLeuIleLeuThrPheAlaIleuSerValThrAspPro 20
Db 138 ATGGGAGGAAGTATCTGTATCTTGAATCTTGACCTTGCCCTCTCTGACAAATCCC 197
QY 21 LeuHisGluLeuLYsAlaIaAlaIaPheProGlnThrThrGluLYsIleSerProAsnTTP 40
Db 198 CTTCAATGAACATAAAGCAAGCTCTTCCCCAGACACATGAGAAATTAAGTCGAATTGG 257
QY 41 GusaRgLYIleAsnValaAspLeuAlaIleSerThrArgGlnThyThtIleuGlnGlnIleu 60
Db 258 GAATTCGGATTAAATGTTGACTTGGCAATTTCCACAGGCAATATATATCTACACAGCTT 317
QY 61 PheTYARgTYGlyGluAsnAsnSerLeuSerValGluGlyPheArgLYsLeuLeuGln 80
Db 318 TTCTACCCCTATGAGGAATAATATCTTGTCTAGTTGAAGGGTTCAGAAATTAATTCA 377
QY 81 AsnIleGlyIleAspLYsIleuLYsArgIleHisIleHisIleAspHisAspHisIleSer 1000

```

Db 378 AATATAGCATAGATTAAGATTTAAAGAAATCCATATACCATGACCGACCATCCTCA 437
 QY 101 AAPHISGIUHIASHISERASPHISGIUARGHISERASPHISGIUHIASHISERASP--- 119
 Db 438 GACCACGAGCATCACTGACGACCATGAGCGTCACTGACGACCATGAGCATCACTGACGAC 497
 QY 119 ----- 119
 Db 498 GAGCATCACTGACCATGATCATCTGACCATTAATCATGCTGCTTGTAATAAT 557
 QY 119 ----- 119
 Db 558 AAGGAAAAAGCTCTTGCCGACGACCATGACTGACTGACTGACTGACTGACTGACTGACT 617
 QY 119 ----- 119
 Db 618 AGCCAGGGGAAAGAGCTCACCGACCGAAGCATGCGGAGTGAAGAAATGTCAGAGAC 677
 QY 119 ----- 119
 Db 678 AGTGTAGTGTAGTGAAGTGAACCTCAACTGTGTCAACAACCTGTCTGAAGAACTGAC 737
 QY 119 ----- 119
 Db 738 TTTCTAGAGACAATAGAGACTCCAAAGACTGGAAAACTCTTCCCAAAAGATGTAGACAGC 797
 QY 119 ----- 119
 Db 798 TCCACTCCACCAGTGTACATCAAAAGAGCGGGTGAAGCGGCTGGTGTAGGAAAAACA 857
 QY 119 ----- 119
 Db 858 AATGAATCTGTAGTGAAGCCCGAAAAAGGCTTTATGTATTCGAAACAACAATGAAAT 917
 QY 119 ----- 119
 Db 918 CCTCAGAGTGTTCATGATCATCAAAAGCTACTGACATCTCATGAGCATGGGCACTCAGGTT 977
 QY 119 ----- 119
 Db 978 CCGGTGAATGCAAGAGATTCAACTATCTGTCCAGCATCATCAACCAATGATGCT 1037
 QY 119 ----- 119
 Db 1038 AGATCTTGTCTGATTCATCAAGTGAAGAAAGGCTGAATCCCTCCAAAGACTTATCA 1097
 QY 119 ----- 119
 Db 1098 TTACAATAGCCTGGGTGGTGTGTTTATAGCCATTTCCATCATGATGATTCCTGTCTG 1157
 QY 119 ----- 119
 Db 1158 CTGGGGGTATCTTAGTGCCCTCTCATGAATCGGGTGTGTTTCAAAATTTCTCTGAGTTTC 1217
 QY 119 ----- 119
 Db 1218 CTGTGGGACGTGCGCGTGGAGCTTTGAGTGTGATGCTTTTTTACACCTTCTTCACAT 1277
 QY 119 ----- 119
 Db 1278 TCTCATGCAAGTCAACCATATAGTCACTAGCATGAAGAACGACATGAAATGAAAAAGA 1337
 QY 119 ----- 119
 Db 1338 GGAGCACTTTGAGTCACTGTCTTCTCAAAAATAGAGAAAGTCTATTTGATTC 1397
 QY 119 ----- 119
 Db 1398 AGGTGAGAGGTCTTAACAGCTTAGAGAGCGCTGTATTTGATGTTCTTGTGAACATGTC 1457
 QY 119 ----- 119
 Db 1458 CTCACATTGATCAACATTTTAAAGTAAAGAAAGAAATCAAGAAAGAACTGAAAT 1517

QY 119 ----- 119
 Db 1518 GATGATGATGATGATTAAGAAAGAGTGTCCAAAGTATGATCTCAACTTTCAACAAT 1577
 QY 119 ----- 119
 Db 1578 GAGAGAAAGTATGATACATGATGATCGAATGAAAGCTATTTTACAGCAGACTCAAGAG 1637
 QY 119 ----- 119
 Db 1638 CCTCCACTTGTATTTCTCAGACGCTGAGCTTTGAGAGAAAGAGTCAATGATGCT 1697
 QY 119 ----- 119
 Db 1698 CAGTCTATCAAGAGAGTCTACATGAATATGTACCCAGAGGCTCAAGATTAATTC 1757
 QY 119 ----- 119
 Db 1758 CATTCACTTTCCAGATACATGCTGGCAGTCAAGAGATCTCATTCAACCATCATGAC 1817
 QY 120 ----- 119
 Db 1818 TACCATCATATTTCTCATCATCAACCAACCAACCAACCATCTCCACAGTCAAGCAGCAG 1877
 QY 128 ArgTyrSerArgGluGluLeuValAspAlaGlyValAlaThrLeuAlaTrpMetValIle 147
 Db 1878 CGTACTCTGGGAGAGCTGAAGAGTGGCGGCTGGCAGCTTGGCTGGATGTGATA 1937
 QY 148 MetGlyAspGlyLeuHisAsnPheSerAspGlyLeuAlaIleGlyValAlaPheThrGlu 167
 Db 1938 ATGGGTATGGCCCTGCACCAATTTTCAAGCATGGCTTACAAATGTGTGCTTTTACTGAA 1997
 QY 168 GlyLeuSerSerGlyLeuSerThrSerValAlaValPheCysHisGluLeuProHisGlu 187
 Db 1998 GGCCTTATCAAGTGTGTTAAGTACTTCTGTGCTGTGTCTGTCTGTCTGTCTGTCTGAA 2057
 QY 188 LeuGlyAspPheAlaValIleLeuValAlaGlyMetThrValIleValAlaValLeuTyr 207
 Db 2058 TTAGGTGCTTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 2117
 QY 208 AsnAlaLeuSerAlaMetLeuAlaTyrLeuGlyMetAlaThrGlyIlePheIleGlyHis 227
 Db 2118 AATGATTTGACGACCAAGCTGGCTATCTTGAATGGCAACAGAAATTTTCTGATGCT 2177
 QY 228 TyrAlaGluAsnValSerMetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisVal 247
 Db 2178 TATGCTGAATAATGTTTCTATGTGATATTTGACCTTCTGCTGCTTATTCATGATGTT 2237
 QY 248 AlaLeuValAspMetValProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSer 267
 Db 2238 GCTCTGTTGATATGATGCTGAAATGCTGCAATGATGCTGATGACCATGATGTTGCT 2297
 QY 268 ArgTrpGlyTyrPhePheLeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeu 287
 Db 2298 CGCTGGGGATTTCTTTTACAGATGCTGGAGTGTGTTGGGTTTGGATTAATGTTA 2357
 QY 288 LeuIleSerIlePheGluHisIleValIlePheArgIleAsnPhe 302
 Db 2358 CTTATTTCCATATTTGAACATTAATCGTGTGTTGATTAATTTTC 2402
 RESULT 14
 US-10-295-027-878
 ; Sequence 878, Application US/10295027
 ; Publication No. US2003023350A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Afari, Daniel
 ; APPLICANT: Azizi, Natasha
 ; APPLICANT: Ginsberg, Wendy M.
 ; APPLICANT: Glash, Kurt C.
 ; APPLICANT: Glynn, Richard
 ; APPLICANT: Hevezi, Peter A.
 ; APPLICANT: Mack, David H.

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OW protein - nucleic search, using frame_plus_p2n model

Run on: March 19, 2005, 22:52:01 ; Search time 213 Seconds

(without alignments)
2319.979 Million cell updates/sec

Title: US-10-659-004-104

Perfect score: 1599

Sequence: 1 MARKSVILITPALSVTNP.....FGIMLLISFEKIVFRINF 302

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1202784 seqs, 81813359 residues

Total number of hits satisfying chosen parameters: 240568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlh
-Q=/cgn2_1/USPRO.spool/US10659004/runat_14032005_181236_15771/app.query.fasta_1.455
-DB=Issued_Patents_NA -QEMT=fastcap -SUFFIX=rni -MINMATCH=0.1 -LOOFC=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPRM=pct -NOR=ext -HEARSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10659004@cgn2_1.1.69 -runat_14032005_181236_15771 -NCEU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=110 -WARN TIMEOUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Issued_Patents_NA:

1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1356.5	84.8	2268	US-09-642-034-4	Sequence 4, Appli
2	1356.5	84.8	3461	US-09-642-034-1	Sequence 1, Appli
3	959	60.0	1310	US-08-311-023-1	Sequence 1, Appli
4	959	60.0	2404	US-08-311-023-3	Sequence 3, Appli
5	439	27.5	4573	US-09-814-915A-29	Sequence 29, Appli
6	395	24.7	2229	US-09-774-528-283	Sequence 283, App
7	374	23.4	1888	US-09-599-360B-38	Sequence 38, Appli
8	283	17.7	464	US-09-621-976-1342	Sequence 1342, Ap
9	212.5	13.3	633	US-09-270-767-2100	Sequence 2100, Ap
10	212.5	13.3	633	US-09-270-767-17382	Sequence 17382, A
11	209.5	13.1	754	US-09-270-767-1621	Sequence 1621, Ap
12	209.5	13.1	754	US-09-270-767-16903	Sequence 16903, A

13	161	10.1	522	US-09-663-600A-149	Sequence 149, App
14	159	9.9	894	US-09-489-039A-5631	Sequence 5631, Ap
15	156	9.8	1203	US-09-543-681A-1979	Sequence 1979, Ap
16	152	9.5	584	US-09-663-600A-55	Sequence 55, Appli
17	143	8.9	10304	US-09-627-465B-1	Sequence 1, Appli
18	143	8.9	34230	US-09-949-016-12052	Sequence 12052, A
19	143	8.9	128470	US-09-949-016-13765	Sequence 13765, A
20	142	8.9	377	US-08-333-766A-1	Sequence 1, Appli
21	140.5	8.8	825	US-09-711-164-198	Sequence 198, App
22	140.5	8.8	825	US-09-711-164-236	Sequence 236, App
23	140.5	8.8	1340	US-09-513-999C-14928	Sequence 14928, A
24	140.5	8.8	1771	US-09-949-016-4307	Sequence 4307, Ap
25	137	8.6	789	US-09-902-540-8863	Sequence 8863, Ap
26	137	8.6	8540	US-09-902-540-944	Sequence 944, App
27	137	8.6	767677	US-09-949-016-12147	Sequence 12147, A
28	137	8.6	767677	US-09-949-016-17361	Sequence 17361, A
29	136	8.5	315	US-09-489-039A-6385	Sequence 6385, Ap
30	135.5	8.5	97195	US-09-949-016-12212	Sequence 12212, A
31	135.5	8.5	97196	US-09-949-016-16971	Sequence 16971, A
32	133.5	8.3	765	US-09-902-540-8591	Sequence 8591, Ap
33	133.5	8.3	6682	US-09-902-540-906	Sequence 906, App
34	133.5	8.3	35154	US-09-949-016-12384	Sequence 12384, A
35	133.5	8.3	39154	US-09-949-016-12801	Sequence 12801, A
36	133.5	8.3	39443	US-09-949-016-14326	Sequence 14326, A
37	133.5	8.3	39443	US-09-949-016-14327	Sequence 14327, A
38	133.5	8.3	64610	US-09-949-016-12214	Sequence 12214, A
39	132.5	8.3	1769	US-09-428-636-3	Sequence 3, Appli
40	132	8.3	30656	US-09-949-016-14613	Sequence 14613, A
41	131	8.2	492	US-09-902-540-6150	Sequence 6150, Ap
42	131	8.2	1479	US-09-902-540-297	Sequence 297, App
43	131	8.2	27933	US-09-949-016-12369	Sequence 12369, A
44	131	8.2	142783	US-09-949-016-15127	Sequence 15127, A
45	131	8.2	767677	US-09-949-016-12147	Sequence 12147, A

ALIGNMENTS

RESULT 1
US-09-642-034-4
Sequence 4, Application US/09642034
Patent No. 6762020
GENERAL INFORMATION:
APPLICANT: Mack, David
APPLICANT: Gish, Kurt C.
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: No. 6762020el Methods of Diagnosing Breast Cancer,
TITLE OF INVENTION: Compositions, and Methods of Screening for Breast
FILE REFERENCE: 018501-009700US
CURRENT APPLICATION NUMBER: US/09/642,034
CURRENT FILING DATE: 2000-08-18
PRIOR APPLICATION NUMBER: US 09/268,865
PRIOR FILING DATE: 1999-03-15
PRIOR APPLICATION NUMBER: US 09/450,810
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: US 09/453,137
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: US 09/525,361
PRIOR FILING DATE: 2000-03-15
PRIOR APPLICATION NUMBER: WO PCT/US00/06952
PRIOR FILING DATE: 2000-03-15
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 2268
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: open reading frame encoding human breast cancer
US-09-642-034-4
Alignment Scores:


```

Db      2041  TATGCTGAAAAGTTTCTATGAGAAATTTGCACTTACGCTGCTTATTCATGATGTT 1100
Qy      248   ALaLeuValAspMetValProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSer 267
Db      2101  GCTCGGTGATATGTTACTGTAAGAAAGCTGCACAAATGATCTGATGACCATGATGACG 216
Qy      268   ArgTTPGlyTYrPhePheLeuGlnAanaAlaGlyMetLeuLeuGlyPheGlyIleMetLeu 287
Db      2161  CGCTGGGGGTATTTCTTTTACAGAAATGCTGGCGATCTTTGGGTTTGGAAATTAATGTA 222
Qy      288   LeuIleSerIlePheGlnHisLysIleValPheArgIleAsnIlePhe 302
Db      2221  CTTATTTCCATATTTGAACATAAATGCTGTTTGTAATAATTTTC 2265

RESULT 2
US-09-642-034-1
; Sequence 1, Application US/09642034
; Patent No. 6762020
; GENERAL INFORMATION:
; APPLICANT: Mack, David
; APPLICANT: Gish, Kurt C.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: No. 6762020el Methode of Diagnosing Breast Cancer,
; TITLE OF INVENTION: Compositions, and Methods of Screening for Breast
; TITLE OF INVENTION: Cancer Modulators
; FILE REFERENCE: 018501-009700US
; CURRENT APPLICATION NUMBER: US/09/642, 034
; CURRENT FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 09/268,865
; PRIOR FILING DATE: 1999-03-15
; PRIOR APPLICATION NUMBER: US 09/450,810
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: US 09/453,137
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: US 09/525,361
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: WO PCT/US00/06952
; PRIOR FILING DATE: 2000-03-15
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3461
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human breast cancer protein BCR4 CDNA
; NAME/KEY: CDS
; LOCATION: (138)..(2405)
; OTHER INFORMATION: human BCR4
US-09-642-034-1

Alignment Scores:
Pred. No.:          9,24e-148          Length:          3461
Score: No.:         1356.50           Matches:          301
Percent Similarity: 40.00%           Conservative:      0
Best Local Similarity: 39.87%         Mismatches:       1
Query Match:        84.83%           Indels:           453
DB:                 4                 Gaps:             1

US-10-659-004-104 (1-302) x US-09-642-034-1 (1-3461)
Qy      1     MetaLaArgLysLeuSerValIleLeuIleLeuThPheAlaLeuSerValThraAspPro 20
Db      138   ATGGGAGGAAAGTTATCTGTATCTGTATCTGTACCTTGGCCCTCTGTGCACAAATCCC 197
Qy      21     LeuHISGluLeuLysAlaAlaPheProGlnThrThrGluLysIleSerProAsnTrp 40
Db      198   CTTACATGAACATAAAGCAGCTGCTTCCGCCACAGCACATGAGAAATTAATTCGGAATTGG 25
Qy      41     GUGSERGlyIleAsnValAspLeuAlaIleSerThrArgGlnTyrrHisLeuGlnGlnLeu 60

```

Db	258	GAATTCGGCATTTAATGTGACTTGCGCAATTTCCACAGCGCAATATCATCTCAACACAGTT	317
Qy	61	PheTyrArgTyrGlyGluAsnAsnSerLeuSerValGlyIlePheArgLysLeuLeuGln	80
Db	318	TTCTAACCGCTATGGAGAAATATATCTTCTGTCAGTTGAAGGGTTCAAGAAATTACTTCA	377
Qy	81	AsnIleGlyIleAspLysIleLysArgIleHisIleHisHisAspHisAspHisHisSer	100
Db	378	AATATAGCGCATAGATTAAGATTAAAGATTCATATACCCATGACACAGACCATCATCTCA	437
Qy	101	AspHisGlyHisHisSerAspHisGlyLysHisSerAspHisGlyHisHisHisSerAsp	119
Db	438	GAACACAGCATCTACTCAGACCAAGAGCGTCACTCAGACCATGACATCATCTCAGACCA	497
Qy	119	-----	119
Db	498	GAGCATCATCTGACCATGATCATCATCTCACCATATACGTGCTTCTGTGTAATAAT	557
Qy	119	-----	119
Db	558	AAGCGAAAAGCTCTTTGCCAGACCATGATCTAGATAGTTCAAGTCTAGAAAC	617
Qy	119	-----	119
Db	618	AGCCAGGGGAAAAGAGCTCACCGACACGACATGCGCAGTGTAGAGAAATGTCAAGAC	677
Qy	119	-----	119
Db	678	AGTGTAGTGTAGTGAAGTGAAGTCACTCAACTGTGTGAACAACACTGTCTCGAAGGAAC	737
Qy	119	-----	119
Db	738	TTTCTAGAGACAATAGAGACTCCAAAGCTGAAAACTTCCCCAAAGATGTAAAGAC	797
Qy	119	-----	119
Db	798	TCCACTCCACCAGTGTCACTCAAAGACCGGGTGAAGCGGCTGTGTAAGAAAAACA	857
Qy	119	-----	119
Db	858	AATGAATCTGAGTGAAGGACCCCGAAGAGCTTTATGTATTCACAGAAACACAATGAAAT	917
Qy	119	-----	119
Db	918	CCTCAGAGTGTTCATATGATCATCAAGCTACTGACATCTCATGCGATGGCATCCAGGTT	977
Qy	119	-----	119
Db	978	CCGCTGAATGCACAGAGAGTTCACATCTCTGTCCAGGCATCATCAACCAATTGAGCT	1037
Qy	119	-----	119
Db	1038	AGATCTGTCTGATTCATACCAAGTGAAGAAAGAGCTGAATCCCTCCAAAGACCTATCA	1097
Qy	119	-----	119
Db	1098	TTACAAATAGCCGTGGGTGTGTGTTTATAGCCATTTCCATCATCATGTTCTCTGTCTG	1157
Qy	119	-----	119
Db	1158	CTGGGGGTTATCTTAGTGCCTTCATGAAATCGGGTGTTCAAATTTCTCTGAGTTTC	1217
Qy	119	-----	119
Db	1218	CTTGAGCACTGGCGGTGGGACTTTGAGTGTATGCTTTTAAACCTTCTTCACAT	1277
Qy	119	-----	119
Db	1278	TCTCATGAGTACACACCATATGTCATGACATGAAGAACCAGCAATGGAATGAAGAAGA	1337
Qy	119	-----	119
Db	1338	GGACCACTTTTCACTCATCTGTCTTTCAAAAATAGAGAAGAGTGCCTAATTTTGATTC	1397

OY	119	-----	119
Db	1398	ACGTGAAGGGCTTAAACAGCTTAGAGAGCGCTGATATTTCAATGTTCTTGTGGAACATGTC	1457
OY	119	-----	119
Db	1458	CTCAGATTGATCAACAATTTAAAGATTAAGAGAAAAAAGATCAGAGAAACCTGAAAAT	1517
OY	119	-----	119
Db	1518	GATGATGATGTGGAGATTAAAGACAGTTGTTCACAGTATGAAATCTCAACTTCAACAAT	1577
OY	119	-----	119
Db	1578	GAGGAGAAAGTAGATACAGATGATGACACTGAAGGCTATTTACGACGAGACTCCAAAG	1637
OY	119	-----	119
Db	1638	CCCTCCACTTTGATTCTCAGCAGCTTGCACTTGGAAGAAAGAGCTCATGATAGCT	1697
OY	119	-----	119
Db	1698	CATGCTCATCCACAGGAAGTTCACATGAATATGTACCAGAGGGTGCAAGATTAATGC	1757
OY	119	-----	119
Db	1758	CATTGACATTTCCACGATACACTCGGCCAGTACGACATCTCATTCACCACCATATGAC	1817
OY	120	-----	120
Db	1818	TACCATCATATTTCTCCATCATGACCAACCAACCAACCAACCACTCCACAGTCACGCGCAG	1877
OY	128	AGCTTTSerArgGluGluLeuIysAspAlaValAlaThrLeuAlaTrpMetValIle	147
Db	1878	CGCTACTCTCGGAGAGCTGAAAGATGCGCGGGTCCGACCTTGGCTGGATGGTGATG	1937
OY	148	MetGlyAspGlyLeuHisAsnPheserAspGlyLeuAlaIleGlyAlaAlaPheThrGlu	167
Db	1938	ATGGGGATGAGGCTCGCACATTTCAACGATGGCCGTGACATTTGGTCTTCTACTGAA	1997
OY	168	GlyLeuSerSerGlyLeuSerThrSerValAlaValPheCysHisGlyLeuProHisGlu	187
Db	1998	GGCTTATCAAGAGGTGTTAAGTACTTCTGTGCTGTGTCTGTGCATAGTTGCTCATGA	2057
OY	188	LeuGlyAspPheAlaValLeuLeuLysAlaGlyMetThrValLysGlnAlaValLeuIyr	207
Db	2058	TTAGTGACCTTGGCTGTCTGTACTAAAGGCTGGCATGACCGTAAAGCAGGCTGCTTTAT	2117
OY	208	AsnAlaLeuSerAlaMetLeuAlaTrpLeuGlyMetAlaThrGlyIlePheIleGlyHis	227
Db	2118	AATGCAATTTGCACCCATGCTGGGGATCTTGGAGATGCAACAGAAATTTTCATTGGTCA	217
OY	228	TyrAlaGluAsnValSerMetCripIlePheAlaLeuThrAlaGlyLeuPheMetHisVal	247
Db	2178	TATGCTGAAATGTTTCTATGTGATATTTGGCACTTACTGCTGGCTTATTCATGATGTG	2237
OY	248	AlaLeuValAspMetValProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSer	267
Db	2238	GCCTCGATTATATGTGTAACCTGAAAGCTCACAAATGATCTGATGCAATGGATGAC	2297
OY	268	ArgTrpGlyTyrPhePheLeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeu	287
Db	2298	CGCTGGGGGATATTTCTTTTACGAAGATGCTGGAGCTTTGGGTGTTTGGAAATTAATG	2357
OY	288	LeuIleSerIlePheGluHisIleValPheArgIleAsnPhe	302
Db	2358	CTTATTTTCATATTTGAACATAAATGCTGTGTTGATTAATTTTC	2402

RESULT 3
US-08-311-023-1
; Sequence 1, Application US/083110233
; Patent No. 5693465

```

GENERAL INFORMATION:
APPLICANT: MANNING, David Lockwood
APPLICANT: NICHOLSON, Robert Ian
APPLICANT: GEE, Julia Margaret
APPLICANT: GREEN, Christopher Douglas
TITLE OF INVENTION: METHODS FOR PREDICTING THE BEHAVIOUR OF
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Young & Thompson
STREET: 745 South 23rd Street
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DO5
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/311,023
FILING DATE: 22-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: PATCH, Andrew J.
REGISTRATION NUMBER: Reg. No. 5693465 32,925
REFERENCE/DOCKET NUMBER: WCM.56
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703/521-2297
TELEFAX: 703/685-0573
TELEX: 248425

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1310 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1293
US-08-311-023-1

Alignment Scores:
Pred. No.: 5,32e-102 Length: 1310
Score: 959.00 Matches: 194
Percent Similarity: 90.41% Conservative: 4
Best Local Similarity: 86.58% Mismatches: 19
Query Match: 59.97% Indels: 5
DB: 1 Gaps: 1

US-10-659-004-104 (1-302) x US-08-311-023-1 (1-1310)
QY      86 LyslllelysaATlIehlgllshlsyaerPhlsAerPhls-----Glu 103
       ::::| | | | | | | | | | | | | | | | | | | | | | | | | |
Db      629 ACAAATAAATGCCATTCACATTTTC-CAGCATCACTCGCCAGTCAAGCATTCATTCAC 687
       ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      104 HlsHlsesAsrPhlsGLuLrgHlsesAsrPhlsGLUHLHlsesArPhlsHlsProHis 123
       ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      688 CACCATCATGCATCACATCATATATTTCTCCATCATACACAACAACACCATCTTCAC 747
       ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      124 SerHlsesGrlnArGTySerArgrglULeuLyAsrPalaglYValatrrLeuAla 143
       ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      748 AGTCACAGCAGGACTCTCTGGGAGGAGCTGAAGAATGCCGGCGCTGGCATTTGGCC 807
       ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      144 TrPMeValIlleMcGlVasrPlYLeuHlsanPrhesAsrPlYLeuAlallegLYala 163
       ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      808 TGcATgATGATATAgGGtATGgcCTGCACAATTTCAAGcATggcCTTAGcAAttGGTGCT 867
       ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      164 AlaphetHrGlUGlYLeusersErGlYLeuSerrThSerVaLaLavAlpheCyHlsIGLU 183
       ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      868 GCttttTAAGAGcctTtAcAtGtgTtTAAGrActTtcgtTcgTgTtGtGCATGAG 927
       ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Qy 184 LeuprohisguleuglyaspphealavaluenuylsaglmetThVallyegln 203
Db 928 TTGCTCATGATTAAGTGGACTTGTCTTACTTAAGGCTGGCATGACCGTTAAGCAG 987
Qy 204 AlavaluenuylsagulaaleuSeralametuallaryleuglymetlathrGlylle 223
Db 988 GCTGCTCTTATATATGATTCATGACCATGCTGGGATCTTGGATGGAACAGGATTT 1047
Qy 224 PheilegylhistryalaglualsvalsermetrllpheaaleuthrAlaglyleu 243
Db 1048 TTCATGTCATATGCGAAGAAATGTTCTATGTGATATTTGCATCTGCTGGCTTA 1107
Qy 244 PheethsvalalaleuvalaspmetvalproglumetuhsaanaaspaSersp 263
Db 1108 TTCATGATGTCCTGCTGATGATGATGACCTGCAAAATGTCACAAATGATGATGAC 1167
Qy 264 HisgilycysSeraGTPGlyTyrrPhepheuglnasnaaglmetleuGlyPhe 283
Db 1168 CATGATGTAAGCCCTGGGGATATTTCTTTTACAGATGCTGGGATCTTTGGGTTT 1227
Qy 284 GlyllemetleuenuleserllephegluhlsylsilevalPheargileasnphe 302
Db 1228 GGAATTAATGTAATTAAT-TCCATATTTGAACATTAATCGTTCGT-ATAAATTC 1282

RESULT 4

US-08-311-023-3
Sequence 3, Application US/08311023
Patent No. 5693465
GENERAL INFORMATION:
APPLICANT: MANNING, David Lockwood
APPLICANT: NICHOLSON, Robert Ian
APPLICANT: GEE, Julia Margaret
APPLICANT: GREEN, Christopher Douglas
TITLE OF INVENTION: METHODS FOR PREDICTING THE BEHAVIOUR OF
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Young & Thompson
STREET: 745 South 23rd Street
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/311,023
FILING DATE: 22-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: PATCH, Andrew J.
REGISTRATION NUMBER: Reg. No. 5693465 32,925
REFERENCE/DOCKET NUMBER: WCM.56
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703/521-2297
TELEFAX: 703/685-0573
TELEX: 248425
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2404 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-08-311-023-3

Alignment Scores:
Pred. No.: 1.37e-101 Length: 2404
Score: 959.00 Matches: 194
Percent Similarity: 90.41% Conservative: 4
Best Local Similarity: 88.58% Mismatches: 19

Query Match: 59.97% Indels: 5
DB: 1 Gaps: 1
US-10-659-004-104 (1-302) x US-08-311-023-3 (1-2404)

Qy 86 LysileysarglIenhlsienhlsasphlsasphlsisSeraaphls-----Glu 103
Db 629 AGAATTAATGCAATTCATTC-CACGATACACTCGGCCAGTACAGATCTCATTCAC 687
Qy 104 HislsSeraphlsaglualsvalsermetrllpheaaleuthrAlaglyleu 123
Db 688 CACCATATGACTACATCATATTCATCATCACACCAACCAACCATCTCAC 747
Qy 124 SerlsSerglnaGlyTyrrPhepheuglnasnaaglmetleuGlyValAlaThleuAla 143
Db 748 AGTACAGCCAGCCCTCTCGGAGAGAGCTGAAGATGGCGGCTGCCCATTTGGCC 807
Qy 144 TrpmetValIlemetGlyaspglyleuhsaanaaspaSeraGlyleuAlaIleGlyAla 163
Db 808 TGGATGATGATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 867
Qy 164 AlaphethrgluglyleuSeraGlyleuSeraThservalalavalPheCysHisglu 183
Db 868 GCTTTACTGAAGCCTTATATCAAGTGTTAAGTACTCTGCTGCTGCTGCTGCTGCTGCTG 927
Qy 184 LeuprohisguleuglyaspphealavaluenuylsaglmetThVallyegln 203
Db 928 TTGCTCATGATTAAGTGGACTTGTCTTACTTAAGGCTGGCATGACCGTTAAGCAG 987
Qy 204 AlavaluenuylsagulaaleuSeralametuallaryleuglymetlathrGlylle 223
Db 988 GCTGCTCTTATATGATTCATGACCATGCTGGCGATCTTGGATGGAACAGGATTT 1047
Qy 224 PheilegylhistryalaglualsvalsermetrllpheaaleuthrAlaglyleu 243
Db 1048 TTCATGTCATATGCGAAGAAATGTTCTATGTGATATTTGCATCTGCTGGCTTA 1107
Qy 244 PheethsvalalaleuvalaspmetvalproglumetuhsaanaaspaSersp 263
Db 1108 TTCATGATGTCCTGCTGATGATGATGACCTGCAAAATGTCACAAATGATGATGAC 1167
Qy 264 HisgilycysSeraGTPGlyTyrrPhepheuglnasnaaglmetleuGlyPhe 283
Db 1168 CATGATGTAAGCCCTGGGGATATTTCTTTTACAGATGCTGGGATCTTTGGGTTT 1227
Qy 284 GlyllemetleuenuleserllephegluhlsylsilevalPheargileasnphe 302
Db 1228 GGAATTAATGTAATTAAT-TCCATATTTGAACATTAATCGTTCGT-ATAAATTC 1282

RESULT 5

US-09-814-915A-29
Sequence 29, Application US/09814915A
Patent No. 6750015
GENERAL INFORMATION:
APPLICANT: Horwitz, Kathryn
APPLICANT: Richer, Jennifer
TITLE OF INVENTION: Progesterone Receptor-Regulated Gene Expression and Methods Relat
FILE REFERENCE: 2848-39
CURRENT APPLICATION NUMBER: US/09/814,915A
PRIOR FILING DATE: 2002-03-21
CURRENT APPLICATION NUMBER: 60/214,870
PRIOR FILING DATE: 2000-06-28
NUMBER OF SEQ ID NOS: 108
SOFTWARE: Patentin version 3.1
SEQ ID NO 29
LENGTH: 4573
TYPE: DNA
ORGANISM: Homo sapiens
US-09-814-915A-29

Alignment Scores:
Pred. No.: 1.96e-40 Length: 4573


```

Db      302 CTGCTTACGCCCCAT-----GTGCTTCTTACAGGGGCTGGGCTG 340
      280 LeuGluGlyPheGlyIleMetLeuLeuIleSerIlePheGlyHisIleVal 297
      341 CTGCTGGGGGGGGCGGCTCATCTTGCCTTAACCCCTGCGAGGAGCGGCTACTG 394

RESULT 9
US-09-270-767-2100
; Sequence 2100, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2100
; LENGTH: 633
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-2100

Alignment Scores:
Pred. No.:      2,54e-15      Length:      633
Score:          212.50      Matches:      47
Percent Similarity: 65.22%      Conservative: 28
Best Local Similarity: 40.87%      Mismatches: 29
Query Match:    13.29%      Indels:      11
DB:             4           Gaps:          4

US-10-659-004-104 (1-302) x US-09-270-767-2100 (1-633)

QY      187 GluLeuGlyAspPheAlaValLeuLeuIleGlyMetThrValIleGlnAlaValLeu 206
      3 GAACCTGGGGGACCTTGTCTGTCTGTCTCCAAACAGGTGTCTCATCGAAGACCCCTCAT 62
      207 TyrAsnAlaLeuSerAlaMetLeuLeuIleGlyMetAlaThrGlyIlePheIleGly 226
      63 ATGAACATTGTAGCTCAAGTCTTGAAGCTTGTGGGCGATGTCGGTGGCTGTATTGCT 122
      227 HisTyrAlaGluAsnValSerMetTrrIlePheAlaLeuThrAlaGlyLeuPheMetHis 246
      123 GGAATTGGAGACGGCATGACCCCAATGATTTATGACGCCGCGTGTCTTCTCTGAT 182
      247 ValAlaLeuValAspMetValProGluMet-----LeuHisAsnAsp-----AlaSer 262
      183 ATTGCTTTTGGCGATCTGTGTCGCCACAAATGATGTGGCCCAATCCGGAATGGCCAA 242
      263 Asp---HisGlyCysSerArgTrrGlyTyrPhePheLeuGlnAsnAlaGlyMetLeuLeu 281
      243 GATCCAAAGGCT-----ATTAATATACAAATTGTTGTATCTCTTG 284
      282 GlyPheGlyIleMetLeuLeuIleSerIlePheGlyHisIleVal 296
      285 GGTGAGCTGATATATGTGGCTATTTGCCCTCAATGAACAGCATTTA 329

RESULT 10
US-09-270-767-17382
; Sequence 17382, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 17382
; LENGTH: 633
; TYPE: DNA

```

```

; ORGANISM: Drosophila melanogaster
US-09-270-767-17382

Alignment Scores:
Pred. No.:      2,54e-15      Length:      633
Score:          212.50      Matches:      47
Percent Similarity: 65.22%      Conservative: 28
Best Local Similarity: 40.87%      Mismatches: 29
Query Match:    13.29%      Indels:      11
DB:             4           Gaps:          4

US-10-659-004-104 (1-302) x US-09-270-767-17382 (1-633)

QY      187 GluLeuGlyAspPheAlaValLeuLeuIleGlyMetThrValIleGlnAlaValLeu 206
      3 GAACCTGGGGGACCTTGTCTGTCTGTCTCCAAACAGGTGTCTCATCGAAGACCCCTCAT 62
      207 TyrAsnAlaLeuSerAlaMetLeuLeuIleGlyMetAlaThrGlyIlePheIleGly 226
      63 ATGAACATTGTAGCTCAAGTCTTGAAGCTTGTGGGCGATGTCGGTGGCTGTATTGCT 122
      227 HisTyrAlaGluAsnValSerMetTrrIlePheAlaLeuThrAlaGlyLeuPheMetHis 246
      123 GGAATTGGAGACGGCATGACCCCAATGATTTATGACGCCGCGTGTCTTCTCTGAT 182
      247 ValAlaLeuValAspMetValProGluMet-----LeuHisAsnAsp-----AlaSer 262
      183 ATTGCTTTTGGCGATCTGTGTCGCCACAAATGATGTGGCCCAATCCGGAATGGCCAA 242
      263 Asp---HisGlyCysSerArgTrrGlyTyrPhePheLeuGlnAsnAlaGlyMetLeuLeu 281
      243 GATCCAAAGGCT-----ATTAATATACAAATTGTTGTATCTCTTG 284
      282 GlyPheGlyIleMetLeuLeuIleSerIlePheGlyHisIleVal 296
      285 GGTGAGCTGATATATGTGGCTATTTGCCCTCAATGAACAGCATTTA 329

RESULT 11
US-09-270-767-1621
; Sequence 1621, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1621
; LENGTH: 754
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-1621

Alignment Scores:
Pred. No.:      7,46e-15      Length:      754
Score:          209.50      Matches:      46
Percent Similarity: 51.22%      Conservative: 38
Best Local Similarity: 28.05%      Mismatches: 69
Query Match:    13.10%      Indels:      11
DB:             4           Gaps:          3

US-10-659-004-104 (1-302) x US-09-270-767-1621 (1-754)

QY      135 LysAspAlaGlyValAlaThrLeuAlaTrrPheValIleMetGlyAspGlyLeuHisAsn 154
      1 AAGAAACAGCGGACGAGAGGTGGCTGATATGAACCTTGGCCCAACTCAATTTGCAAT 60
      155 PheSerAspGlyLeuAlaIleGlyValAlaPheThrGlyGlyLeuSerSerIleLeuSer 174
      61 TTCACACAGCGTCTGCGCTGTGATCTTTGTGTCTTCAAGACAGCGCATCTTA 120

```



```

Db      445 CACGGAACATAGCCATGAGAC-----ACGAAGAACAATTAGTCACGTAATTGTAATGAC 498
Oy      240  PALAGLEUPHMECHISVALALEUVALMERETVALPROGIMLEUHSIASHNA 260
           :::::|||||:|||||:|||||:|||||:|||||:|||||:
Db      499 AACATGCAATTATGAGGCAAGAGAGACTGGTATCTCATGCTCCCGCAATTAGCCAAAACGA 558
Oy      260 PALASERAPHNIAIGLYCYE-----SERVHTPTGLYTYRPE 272
           |||||:|||||:|||||:|||||:|||||:|||||:
Db      559 A-----TGTAAATAATGATGATGATGATCTG 586

```

Search completed: March 20, 2005, 01:39:54
Job time : 225 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_p1uc_model

Run on: March 19, 2005, 21:58:34 ; Search time 4918 Seconds
(without alignments)
2975.493 Million cell updates/sec

Title: US-10-659-004-104
1599

Perfect score: 1 MARLSTVILITFALSVTNP.....FGIMLLISIFBKIVRINF 302
Sequence:

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=x1h
-Q=/cgm2_1/USPTO.epool/US10659004/runat_14032005_181235_15741/app_query.fasta_1.455
-DB=GenBml -CPMT=faaetap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=b1te -START=1 -END=-1 -MATRIX=biolum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USR=US10659004 @CGN_1_1_3731 @runat_14032005_181235_15741 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenBml:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_on:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1359.5	85.0	2744	6	CQ834422 Sequence
2	1359.5	85.0	2744	6	CQ854099 Sequence
3	1359.5	85.0	2744	9	HSU41060
4	1359.5	85.0	3523	6	CQ493869 Sequence

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	CDS
CQ834422	CQ834422	Sequence 293 from Patent WO2004058805.	CQ834422	CQ834422.1	GI:50833959	Homo sapiens (human)	1	Matsuda, A. and Yoneta, S.	T cell activating gene	Patent: WO 2004058805-A 293 15-JUL-2004;	location/Qualifiers	
											1..2744	
											/organism="Homo sapiens"	
											/mol_type="unassigned DNA"	
											/db_xref="taxon:9606"	
											136..2387	
											/note="unassigned protein product"	
											/codon_start=1	
											/protein_id="CAH05426.1"	
											/db_xref="GI:50833960"	
											/translation="MARKLSVILITFALSVTNPILHLLKAAAPQTTKEKISPMWESGI	
											MDLAISTRQYHLOQLFRRYGENNSLSVEGRKLQNIIGDKIKRIHHHDHSHSH	

ALIGNMENTS

5	1356.5	84.8	3461	6	AX465588
6	1356.5	84.8	3461	6	AX829136
7	1352.5	84.6	3445	6	AX511618
8	1349.5	84.4	2776	6	AX207207
9	1324.5	82.8	2265	6	CQ854111
10	1311.5	82.0	3461	6	AX207205
11	1237.5	77.4	3610	10	BC055012
12	1148	71.8	3287	10	AB071697
13	1007	63.0	2212	9	BC008317
14	1007	63.0	2370	6	CQ834560
15	1002	62.7	2815	10	BC054780
16	959	60.0	1310	6	I76891
17	959	60.0	2404	6	I76892
18	875	54.7	1597	6	BD134432
19	875	54.7	1597	6	AX017261
20	875	54.7	1597	6	AX524965
21	831	52.0	551	6	AX207216
22	817	51.1	2229	5	AB126260
23	740	46.3	1672	9	BC039498
24	718	44.9	3309	6	CQ491400
25	718	44.9	3309	6	CQ497275
26	718	44.9	3760	9	BC073909
27	718	44.9	4417	6	CQ414949
28	718	44.9	5231	9	AB033091
29	718	44.9	5322	6	AX405756
30	712	44.5	3079	10	BC052880
31	712	44.5	3865	10	BC059214
32	712	44.5	5003	10	BC062918
33	709	44.3	4395	10	AK122483
34	674	42.2	3671	5	BC058056
35	663	41.5	160170	2	AP001158
36	659	41.2	221941	2	AC091060
37	659	41.2	224788	2	AP001905
38	634	39.6	736	6	CQ780369
39	634	39.6	736	6	CQ780369
40	634	39.6	736	6	CQ780369
41	634	39.6	736	6	BD125078
42	634	39.6	2863	6	BD125739
43	634	39.6	2863	6	CQ782979
44	634	39.6	2863	6	BD127410
45	594	37.1	152073	2	AC149703

EHHSDEHSHSDHEHSHDHEHSHDHNHAAAGKNNKALCPDHSDSSGKDPNNSGCKA
 HRPEHAGRNVKDSVASAEVTSYVTYVTSVSEGTFLTEIETPRGKXLPKDVSSSTP
 SVTSKRSVRLAGKTNESVSEPRKPMVSNNTNENPOECNASKLITSHMGQVPL
 NATEPNYICPAIINOIDARSLIHTSEKKAIPKTYSLQIAMVGGIATISIFLFL
 LGVILVLEPMNRVFPKFLSLPLVALAVGLSDAFIHLPHSHAHSHSHHEPRAMK
 KRGLPFSHLSSQNI EESAYFDSTWKGALGGLYFMPLVEVHTLTKQFKDKKKRNMK
 KPEMDDVEIKKOLSKYESQSTNEEKVDTDRTGEGYLRADSOEPSHFDDQPAVLE
 EEWMIAMHPOEVNVEYVPRCKNKVSHFDLTGOSDULIHHDHVNHLIHHHQN
 HHPSHSORYSREELKDAGVATLAWVMGDLMFSGDLIGAAIFEGLSGSLSTV
 AVPCHELPHETGDFPAVLKAGMTAKOAVLYNALSAMLYGMATGIFIGHYAEVSWM
 IFALTAGLPMYVALVDVMPKEMHNDASHGSRNGYFPLQAGLILRGIMLISIE
 HKIVRINF"

ORIGIN

Alignment Scores:

Score: 5.5e-112 Length: 2744
 Percent Similarity: 1359.50 Matches: 301
 Best Local Similarity: 40.32% Conservative: 1
 Query Match: 85.02% Mismatches: 0
 DB: 6 Indels: 447
 Gaps: 1

US-10-659-004-104 (1-302) x CQ834422 (1-2744)

QY 1 MetalaarglyleuSerValleuileuThrPheAlaLeuSerValThrAsnPro 20
 DB 138 ATGGGAGGAGATTATCTGTATCTTGATCTTGACCTTGCCCTCTGTCAAAATCCC 197
 QY 21 LeuHleGlyLeuLysAlaAlaAlaPheProGlnThrThrGluLysLeuSerProAsnTrp 40
 DB 198 CTTTCATGAACTAAAGCAGCTCTTCCGCCAGACACATGAGAAATTAATGTCGAAATGG 257
 QY 41 GluSerGlyLeuAsnValAspLeuAlaIleSerThrArgLysIleuGlnLeu 60
 DB 258 GAATCTGGATTAAATTTGACTTGGCAATTTCCACCGCATATCATCAACAGCTT 317
 QY 61 PheTyArgTyGlyGluAsnAsnSerLeuSerValGluGlyPheArgLysLeuGln 80
 DB 318 TTCTACCGCTATGAGAAATTAATCTTGTGCAAGGAGGTTGAGAAATTAATCTTCA 377
 QY 81 AenHleGlyLeuAspLysIleLysArgIleHisIleHisAspHisAspHisIleSer 100
 DB 378 AATATAGGATGATGATTAAGATTAAGATTCATATACCATGACACACATCATCTCA 437
 QY 101 AspHleGluHisIleSerAspHisGluArgHisSerAspHisGluHisIleSerAsp 119
 DB 438 GACACGAGCATCTACTCAGCATGAGCGTCACTGACCATGAGCATCTCAAGCCAC 497
 QY 119 ----- 119
 DB 498 GAGCATCACTGACCATATATCATGCTGCTTGTGTAAATAAGGAAAGCTCTTGGC 557
 QY 119 ----- 119
 DB 558 CCAAGCATGACTCAAGTAGTTAGSTAAAGATCTTAAGAAACGACGAGGGAAGAGCT 617
 QY 119 ----- 119
 DB 618 CACGACCAAGACATGCGAGTGTAGAGGAATGTCAAGACAGTGTAGTGTAGTAA 677
 QY 119 ----- 119
 DB 678 GTGACCTCACTGTGTACAACACTGTCTGTGAAGAACTCACTTGTAGACAAATAGAG 737
 QY 119 ----- 119
 DB 738 ACTCCAGACCTGAAAACTCTTCCCAAGATGTAAAGCAGCTCCACTCCACCCACTGTC 797
 QY 119 ----- 119
 DB 798 ACATCAAGAGCGGGGTGAGCGCGGTGTGTGTAGGAAACAATGAATGTGTGAGTGTAG 857
 QY 119 ----- 119

DB 858 CCCCCAAAGGCTTTATGTTATTCACAGAAACAAATGAATAATCCTCAGAGTGTTCAT 917
 QY 119 ----- 119
 DB 918 GCATCAAGCTACTGACATCTCATGGCAGTGCAGCTCCAGGTGGAATGCAACAGAG 977
 QY 119 ----- 119
 DB 978 TTCAACTATCTCTGTCCAGCATCATCAACCAATTGATGCTAGATCTTGTCTGATTCAT 1037
 QY 119 ----- 119
 DB 1038 ACAAGTAAAGAAAGGCTGAATCCCTCCAAAGACTATTCATTACAAATAGCCTGGGT 1097
 QY 119 ----- 119
 DB 1098 GGTGTTTATAGCATTTCCATCATCAGTTCCGTCTGTGTGGGGTTATCTTAGTG 1157
 QY 119 ----- 119
 DB 1158 CTTCTCATGAATCGGAGTGTTCATTAATTTCTCTGAGTTCTTGTGGCACTGGCGGT 1217
 QY 119 ----- 119
 DB 1218 GGGACTTGAAGTGTGATGCTTTTACACCTTCTTCACATTTCTCATGCAATGACACAC 1277
 QY 119 ----- 119
 DB 1278 CATAGTCATGACCATGAAGAACACGACATGGAATGAAGAGACCACTTTTCAGTCAT 1337
 QY 119 ----- 119
 DB 1338 CTGTCTTCTCAAAACATGAAGAAAGTGCTATTGTTGATTCACGTGAAGGCTTAACA 1397
 QY 119 ----- 119
 DB 1398 GCTCTAGAGGCTGTATTTTCATGTTCTTGTGAACATGCTCTCATGATGATCAACAA 1457
 QY 119 ----- 119
 DB 1458 TTAAAGATTAAGAAAAAGAAATCAGAAAGAACTGAAATGATGATGTGAGATT 1517
 QY 119 ----- 119
 DB 1518 AAGAAAGCTTGTCCAAATGAAATCTCACTTTCACAAATGAGAGAAAGTATACA 1577
 QY 119 ----- 119
 DB 1578 GATGATGAGTGAAGCTATTTACGAGCAGATTCACAAGAGCCTCCACTTGTATCT 1637
 QY 119 ----- 119
 DB 1638 CAGACCTGAGTCTTGGAAGAAAGAGGTGATGATAGTCACTGATTCACACAGAA 1697
 QY 119 ----- 119
 DB 1698 GTTACAAATGATATGTACCAAGAGGTGCAAGATTAATGCATTCATTCATTCACAGAT 1757
 QY 119 ----- 119
 DB 1758 ACATCGGCACTGACAGATCTCATTCACACCATCATGACTACATCATATTCCTCAT 1817
 QY 120 ----- 119
 DB 1818 CATCAACCAACCAAAACCATCTCTCAAGTCAAGCAGCGCTACTCTGGAGAGAG 1877
 QY 134 LeuLysAspAlaGlyValAlaThrLeuAlaTrpMetValIleMetGlyAspGlyLeuHis 153
 DB 1878 CTGAAGATGCCGGGCTGCCACTTTGGCTGAGATGATTAATGGGTGATGGCTGACAC 1937
 QY 154 AenPheSerAspGlyLeuAlaIleGlyAlaAlaPheThrGluGlyLeuSerSerGlyLeu 173

Db 1938 AATTTCAGCGATGCGCTAGCAATTGCTGCTTTACTGAAAGCTTATCAAGTGTTA 1997
Qy 174 SerThiSerValAlaValPheCyhHisGluLeuProHisGluLeuGlyAspPheAlaVal 193
Db 1998 AGTACTCTGTTGCTGTGTTCTGCATGAGTTCCTCCTGATTAAGAGTACCTTCTGTT 2057
Qy 194 LeuLeuValAspAlaGlyMetThrValIlyAspAlaValLeuValAspAlaLeuSerAlaMet 213
Db 2058 CTACTTAAGGCTGGCATGACCGTTAAGCAGGCTGTCTTTATTAATGATGTTCAGCCATG 2117
Qy 214 LeuAlaTyLeuGlyMetAlaThrGlyIlePheIleGlyHisTyAlaGluAsnValSer 213
Db 2118 CTGCGCATCTTGGATGGAAGCAACAGAAATTTTCATGTGTCATTAATGCGAAGATTTTCT 2177
Qy 234 MetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisValAlaLeuValAspMetVal 253
Db 2178 ATGTGGATATTGCACTTACTGCTGGCTTATTCATGTATGTGCTGTGGTGAATATGTA 2237
Qy 254 ProGluMetLeuHisAspAspAlaSerAspHisGlyCysSerArgTrpGlyTyPhePhe 273
Db 2238 CTTAAATGCTGCAATGATGCTAGTGAACCATGAGATGACCGCTGGGGATATTTCTT 2297
Qy 274 LeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeuLeuIleSerIlePheGlu 293
Db 2298 TTACAGATGCTGGAGTCTTTGGGTTTGGAAATTATGTTACTTATTTCCATATTTGAA 2357
Qy 294 HisValIleValPheArgIleAsnPhe 302
Db 2358 CATAAATCGTGTTCGTATTAATTTTC 2384

RESULT 2
LOCUS CO854099 2744 bp DNA linear PAT 23-AUG-2004
DEFINITION Sequence 1 from Patent WO2004067564.
ACCESSION CO854099
VERSION CO854099.1 GI:51510129
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Law, D., Gish, K.C., Murray, R. and Culp, P.
TITLE Compositions against cancer antigen liv-1 and uses thereof
JOURNAL Patent: WO 2004067564-A 1 12-AUG-2004;
PROTEIN DESIGN LABS, INC. (US)
FEATURES
source 1..2744
location/Qualifiers
1..2744 /organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 5-5e-112 Length: 2744
Score: 1359.50 Matches: 301
Percent Similarity: 40.32% Conservative: 1
Best Local Similarity: 40.19% Mismatches: 0
Query Match: 85.02% Indels: 447
DB: 6 Gaps: 1

US-10-659-004-104 (1-302) x CO854099 (1-2744)

Qy 1 MetAlaArgIleLeuSerValIleLeuIleLeuThrPheAlaLeuSerValThrAsnPro 20
Db 138 ATGCGAGAGAAATATCTGTATCTTGATCTTGACCTTCCCTCTGCTGTCAAAATCCC 197
Qy 21 LeuHisGluLeuValAlaAlaPheProGlnThrTrpGlyIleSerProAsnTrp 40
Db 198 CTTATGAACTTAAGCAAGCTGCTTTCCCAAGACCATGAGAAATTAATGCCAATTTGG 257
Qy 41 GluSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTyHisLeuGlnGlnLeu 60

Db 258 GAATCTGGCATTAATGTTGACTTGGCAATTTCACACGGCAATATCATCAACAGCTT 317
Qy 61 PheTyArgTyGlyGluAsnAsnSerLeuSerValGlyIlePheArgIleLeuGln 80
Db 318 TTCTACCGCTATGAGAAATAATTTCTTGTCACTTGAAGGGTTACAGAAATTAATCTTCA 377
Qy 81 AsnIleGlyIleAspIleValArgIleHisIleHisAspHisAspHisHisSer 100
Db 378 AATATAGCATAGATTAAGATTAAAGAAATCCATATACACCATGACACGACCATCACTCA 437
Qy 101 AspHisGluHisHisSerAspHisGluArgHisSerAspHisGluHisHisSerAsp 119
Db 438 GACCAACGACATCACTGACATGACATGAGGCTACCTCAGACCATGACATCACTCAGAC 497
Qy 119 ----- 119
Db 498 GAGCATCACTGACCATTAATCATGCTGCTTCTGTAAATAAGCAAAAGCTTTTGC 557
Qy 119 ----- 119
Db 558 CCAGACATGACTAGATAGTTCAAGTAAGATCTTAAGAAACAGCCAGGGAAGAGCT 617
Qy 119 ----- 119
Db 618 CACCGACCAAGACATGCCAGTGTAGAAAGAAATGTCAAGACAGTGTAGTGTAGAA 677
Qy 119 ----- 119
Db 678 GTGACTCAACTGTGTACAACACTGTCTTGAGAACTCACTTTAGAGACAATAGAG 737
Qy 119 ----- 119
Db 738 ACTCAAGACCTGGAAAACTTTCCCAAGATGTAAAGACTCCATCCACCAAGTTC 797
Qy 119 ----- 119
Db 798 ACATCAAGAGCGGGTGAAGCGGCTGCTGTGTAGAAACAATGAATCTGTAGTAG 857
Qy 119 ----- 119
Db 858 CCCCCAAAGGCTTATGATATTCAGAAACAATAATGAATCTCAGAGTGTTCAT 917
Qy 119 ----- 119
Db 918 GCATCAAGCTACTGACATCTCATGATGGGATCCAGGTTCCGCTGAATGCAACAGAG 977
Qy 119 ----- 119
Db 978 TTCAACTATCTGTCCAGCATCATCAACCAATGTAGTGTATCTTGTCTGATTCAT 1037
Qy 119 ----- 119
Db 1038 ACAAGTAAAGAGGCTGAATCCCTCAAGACTATTCATTACAAATAGCTGGGT 1097
Qy 119 ----- 119
Db 1098 GGGGTTTATAGCATTTTCATCATCATGATTTCTGTCTGTGGGGTTATCTTAGTG 1157
Qy 119 ----- 119
Db 1158 CTTCTATGAATGGGGTGTTCATTAATTTCTCTGAGTTTCTTGTGGACGTGGCGTT 1217
Qy 119 ----- 119
Db 1218 GGGACTTTGAGTGTATGCTTTTACACCTTCTCCATCTCATGCAAGTACACAC 1277
Qy 119 ----- 119
Db 1278 CATATGATAGCATGAGAAACAGCATGAGAAATGAAGAGACCACTTTTCAATCAT 1337
Qy 119 ----- 119
Db 1338 CTGCTTCTCAAAATATAGAAAGAGTGCATTTTGTGATTCAGGTGAAGGGTCTAACA 1397

[illegible]

ACCESSION	complete cds.
VERSION	U01060
KEYWORDS	U01060.2 GI:12711792
SOURCE	
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 2744)
AUTHORS	Green, C. and Morgan, H.
TITLE	Direct Submission
JOURNAL	Submitted (08-FEB-2001) Biochemistry, University of Liverpool, P.O. Box 147, Liverpool L69 3BX, UK 2 (bases 1 to 2744)
REFERENCE	Green, C., Gillholly, E.M. and Walker, N.J. Direct Submission Submitted (21-NOV-1995) Biochemistry, University of Liverpool, P.O. Box 147, Liverpool L69 3BX, UK On Feb 8, 2001 this sequence version replaced gi:1256000.
COMMENT	
FEATURES	Location/Qualifiers
SOURCE	1..2744
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/cell_line="MCF-7 human breast cancer cell line"
	/note="estrogen induced mRNA"
gene	1..2744
	/gene="Liv-1"
CDS	138..2387
	/gene="Liv-1"
	/note="breast cancer"
	/codon_start=1
	/product="estrogen regulated Liv-1 protein"
	/protein_id="AA96258.2"
	/db_xref="GI:12711793"
	/translation="MARKLSTVLTLFALSVTNPLHLKAAFQPTTEKISPNWESGIL NVDAITRQYLOQLFYRGENNLSVEGRKLQIIGDKIRIHINHDHSHD EHSDHRHSOHLSEHSHDNHNAAGNKKALKCPIDSDSGDPNSOGKPD HREPHASGRNRVKDSVASRYTSVTMSRCGTFLTIEMRPBGKLPKXVSSTPE SVTSKRSRLAGRKTINESVBEPKGMYSNTENTNEPQECNASKLISHMGIOVPI NATPENLCPAINIQIDARSLHTSKKAIIPKTYSLQIAWGGLAISIFSLT LGVLPLSNRVFPKFLPSFLVALAVGTLSGDADLHPHSAHHNHSHSEBPAMEX RKGLFSLHSNRIEESAYPESTWKGLTAGLGFMEFLVBEVLTLIKQPKDKKKNO KPPNDDEVIKKOLSKYESOLSTHEEKVDTPDRGYLRADSPSEHDSQPALVE EEHWTAHAHPQEVNYENVPKCCCKSHPFDTGGOSDLIHNNHDYNIHHNNHO HHPHSHSRISREELKQAGVATTAMWTIMDGILNFSFDGLAIGAFTBGLSSGLSTS AVFHCHLPHLEIGDFAVLLKAGMTVKQAVLYNALSAMLAIVDMATGIFIGHYAENVSMIK IFALTACLFMVTVALLVDWPPEMLHNDAHDHGCRWGYFFLQNMGLFGIMLLISIFE HKIVERINF"
ORIGIN	
Alignment Scores:	
Pred. No.:	5.5e-112 Length: 2744
Score:	1359.50 Matches: 301
Percent Similarity:	40.32% Conservative: 1
Best Local Similarity:	40.19% Mismatches: 0
Query Match:	85.02% Indels: 447
Gaps:	9 Gaps: 1
US-10-659-004-104 (1-302) x HSUA1060 (1-2744)	
OY	1 MetaLaagvylELeuservAlIlleuulrPhelaaleuSerValIthrAnPro 20
Db	138 ATGGCGAGGAAGTTATCTGTAATCTTGATCCTGCACCTTCCTCTGTCACAATCCC 197
OY	21 LeuhIsGluleLVeAlAlalapherProgInThrNgIuLyailEsErProAsnTrp 40
Db	138 CTTCATGAACTRAAAGCAGCGCTCTTCCCCCAAGCACATGAGAATAATTAATGCCGAATGG 257
OY	41 GiuserGylileasValaspLeuAlalieserThrargInTyrlHisleuInglnIleu 60
Db	258 GAATTCGGCATTAATGTTGACTTGGCAATTTCCACAGCGCATATATCATCTACACAGCTT 317

Oy 61 PheTYARGTyTgYgluAaPhaSerleuSerValIgluGlyPheArxglYleuLeuGln 80
 Db 318 TTTCACCCCTATGGAGAAATTAATTTCTTGTCAAGTTGAAGGTTCAAGAAATTAATTCA 377
 Oy 81 AenIleGIYIleAspLysIleLysArgIleHisIleHisAspHisAspHisSer 100
 Db 378 AATATAGCATATGATTAAGATTMAAGAAATCCATATACCATGACACAGCATCATCA 437
 Oy 101 AspHisGluHisHisSerAspHisGluHisSerAspHisGluHisSerAsp--- 119
 Db 438 GACCAAGCATCTACTCAGACATGAGCGTCACTCAGACATGACATCATCAGACAC 497
 Oy 119 ----- 119
 Db 498 GAGCATCACTGACCATTAATCATCTGCTTGTGTAATAATAGCAAAAGCTTTTGC 557
 Oy 119 ----- 119
 Db 558 CCAGACATGACTCAGATAGTTCAAGTAAAGATCTAGAAACAGCCAGGGGAAAGAGCT 617
 Oy 119 ----- 119
 Db 618 CACCGACGAAACATGCCAGTGTAGAAAGATGTCAAGACAGTGTAGCTAGTAA 677
 Oy 119 ----- 119
 Db 678 GTGACCTCACTGTGTACAACACTGTCTGTGAAGAACTCACTTTCTAGACAAATAGAG 737
 Oy 119 ----- 119
 Db 738 ACTCCAAACCTGAAAACTTTCCCAAAAGTGAAGAGTCCACTCCACCAGTGTG 797
 Oy 119 ----- 119
 Db 798 ACATCAAAAGCGCGGTAGCCGCTGCTGTAGAAACAAATGAATCTGTAGAGTAG 857
 Oy 119 ----- 119
 Db 858 CCCCCAAAAGGCTTTATGTATTCAGAAAACAAATGAATAATCTCAGAGAGTGTTCAT 917
 Oy 119 ----- 119
 Db 918 GCATCAAAAGCTACTGACATCTCATGCGCATCGGATCCAGGTTCCGCTGAATGCAACAGAG 977
 Oy 119 ----- 119
 Db 978 TTCACTATCTCTGTCCAGCCATCATCAACAAATTGATGCTAGATCTTGTCTGATTCAT 1037
 Oy 119 ----- 119
 Db 1038 ACAAGTAAAGAGAGCTGAAATCCCTCCAAAGACTTATTCATTAACAAATAGCCTGGGTT 1097
 Oy 119 ----- 119
 Db 1098 GGTGTTTTATAGCATTTCCATCATCACTTCTGTCTGTGCGGGGTTATCTTAGTG 1157
 Oy 119 ----- 119
 Db 1158 CCTCTCATGAATCGGGTGTCTTCAAAATTTCTGTAGATTCTGTGCACTGCGGTT 1217
 Oy 119 ----- 119
 Db 1218 GGGACTTGTAGTGTGATGCTTTTTCACCTTTCACATTCATGCAAGTACACAC 1277
 Oy 119 ----- 119
 Db 1278 CATACTCATAGCATGAAGAACACCATGAAATGAAGAAAGACACATTTTCACTCAT 1337
 Oy 119 ----- 119
 Db 1338 CTGTCTCTCAAAACATAGAAAGAGTGCCTATTTTGAATTCACAGTGAAGGGTCAACA 1397
 Oy 119 ----- 119

Db 1398 GCTCTAGAGGCGCTGTATTTCATGTTCTTGTGAACATGCTCACAATGATCAACAA 1457
 Oy 119 ----- 119
 Db 1458 TTTAAAGTAAAGAAAAAAGATCAGAAAGAACTGAAATGATGATGTGAGATT 1517
 Oy 119 ----- 119
 Db 1518 AAGAGCAGTTGTCAGATGAATTCACACTTTCACAAATGAGAGAAAGTAGATACA 1577
 Oy 119 ----- 119
 Db 1578 GATGATGCACTGAAGGCTATTTCAGAGACATCACAAGACCCCTCCACTTTGATTCT 1637
 Oy 119 ----- 119
 Db 1638 CAGAGCCTGCAGTCTTGAAAGAAAGAGTCAATGATAGCTCATGCTCATCCAGAA 1697
 Oy 119 ----- 119
 Db 1698 GTTACAAATGAATATGTAACCAAGAGGTGCAAGAAATAATGCCATTCAATTCACAGAT 1757
 Oy 119 ----- 119
 Db 1758 ACACTGGCCAGTCAGACGATCTATTCCACCATCACTGACTACATATTTCTCAT 1817
 Oy 120 ----- 1817
 Db 1818 CATCACACACCAAAACCATCTCTCAGTCAACAGCCAGCGCTACTCTCGGAGAGAG 1877
 Oy 134 LeuLysAspAlaGlyValAlaThrLeuAlaTrpMetValIleMetGlyAspGlyLeuHis 1937
 Db 1878 CTGAAGATGCGCGGCTGCCACTTGGCTGGATGGATGAATGAGTGAAGCTGAC 1937
 Oy 154 AsnProSerAspGlyLeuAlaIleGlyValAlaPheThrArgIleGlyLeuSerSerGlyLeu 173
 Db 1938 AATTTACAGCATGGCTTACGATGCTGTCTGCTTTTACTGAAGCTTATCAAGTGTTHA 1997
 Oy 174 SerThrSerValAlaValPheCysHisGluLeuProHisGluLeuGlyAspPheAlaVal 193
 Db 1998 AGTACTTCTGTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 2057
 Oy 194 LeuLeuLysAlaGlyMetThrValIleGlnAlaValLeuIleThrAsnAlaLeuSerAlaMet 213
 Db 2058 CTACTAAAGCGCTGCATGACCGTTAAGCAGGCTGCTTTATATGATGATGCTGACGCCATG 2117
 Oy 214 LeuAlaTyLeuGlyMetAlaThrGlyIlePheIleGlyHisTyAlaGluAsnValSer 233
 Db 2118 CTGGGATCTTGGATGGCAACAGAAATTTCAATTGCTATGCTGAAATATGTTCT 2177
 Oy 224 MetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisValAlaLeuValAspMetVal 253
 Db 2218 ATGTGATATTTGCACTTACTGCTGCTGCTTATTCATGATGATGCTGCTGCTGCTGCTGCT 2237
 Oy 254 ProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSerArgTrpGlyTyrPhePhe 273
 Db 2238 CCGAAATGCTGCACAAATGATGCTAGTACCATGATGATGATGATGATGATGATGATGATGAT 2297
 Oy 274 LeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeuLeuIleSerIlePheGln 293
 Db 2298 TTACGAATGCTGGATGCTTTGGGTTTGGAAATATGATGATGATGATGATGATGATGATGAT 2357
 Oy 294 HisLysIleValPheArgIleAsnPhe 302
 Db 2358 CATAAATCGTGTTCGTATAAATTTTC 2384
 RESULT 4
 CQ493869 3523 bp DNA linear PAT 30-JAN-2004
 LOCUS CQ493869
 DEFINITION Sequence 25736 from Patent WO0160860.
 ACCESSION CQ493869
 VERSION CQ493869.1 GI:41459488

KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
 AUTHORS Schlegel, R., Endege, W.O. and Monahan, J.E.
 TITLE Genes differentially expressed in human prostate cancer and their use
 JOURNAL Patent: WO 0160860-A 25736 23-Aug-2001;
 Millennium Predictive Medicine, Inc. (US)

FEATURES
 source
 Location/Qualifiers
 1..3523
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Alignment Scores:
 Pred. No.: 7,68e-112 Length: 3523
 Score: 1359.50 Matches: 301
 Percent Similarity: 40.32% Conservative: 1
 Best Local Similarity: 40.19% Mismatches: 0
 Query Match: 85.02% Indels: 447
 DB: Gaps: 1

US-10-659-004-104 (1-302) x CQ493869 (1-3523)

QY 1 MetAlaArgLysLeuSerValIleLeuIleLeuThrPheAlaLeuSerValThrAsnPro 20
 Db 138 ATGGCAGAAAGTTATCTGTAATCTTGATCTGACCTTGGCCCTCTCTCAATGCC 197

QY 21 LeuHISgluLeuLysAlaAlaPheProGlnThrGluLysIleSerProAsnTrp 40
 Db 198 CTTGATGAACTAAAGCAGCTGCTTCCCCAGACCACTGAGAAATTAATGTCGAATTGG 257

QY 41 GluSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTrpHisLeuGlnLeu 60
 Db 258 GAATCGGATTAATGTTGATCTGGCAATTTCCACGGCGAATATCATCTCAACAGCTT 317

QY 61 PheTyrArgTyrGlyLysAsnAsnSerLeuSerValGluGlyPheArgLysLeuGln 80
 Db 318 TTCTACCGCTATGAGAAATAATTTCTTGTGCTGAGGAGGTTGAGAAATTAATCTCA 377

QY 81 AsnIleGlyIleAspLysIleLysArgIleHisIleHisIleAspHisAspHisSer 100
 Db 378 AATATAGGCAATGATTAAGATTAAAGATCCATATACCATGACCAACCATCACTCA 437

QY 101 AspHisGluHisIleSerAspHisGluArgHisSerAspHisGluHisIleSerAsp 119
 Db 438 GACCAAGAGCATCACTGACATGAGCGTCACTCAAGACATGACATCACTGACAC 497

QY 119 -----
 Db 119 -----

QY 498 GAGCATCACTGACCATATATCATGCTGTTCTGTGTAATAATAAGGAAAGCTTTTGC 557
 Db 119 -----

QY 558 CCAGACCATGACTCAGATAGTTCAGGTAAAGATCTTAAGAACAGCAGGGGAAAGAGCT 617
 Db 119 -----

QY 618 CACCGACCAAGAACGTCAGTGTGAAGAAATGTCAAGACAGTGTAGTGTAGTGA 677
 Db 119 -----

QY 678 GTGACCTCAACTGTGTACAACACTGTCTGGAAGAACTCACTTTAGAGACATAGAG 737
 Db 119 -----

QY 738 ACTCCAGACCTGGAATACTTTCCCAAGATGTAGAGCTCCACTCCACCACTGTC 797
 Db 119 -----

Db 798 ACATCAAGAGCCGGGTGACCGGCTGCTGTAGAGAAACAAATGAATCTGTAGTAGAG 857
 QY 119 -----

Db 858 CCCCCAAAAGCTTTATGTATATCCAGAACCAAAATGAAAATCTCAGAGTGTTCAT 917
 QY 119 -----

Db 918 GCATCAAGCTACTGACATCTCATGCGATGGGATCCAGGTTCCGCTGAATGCAACAGAG 977
 QY 119 -----

Db 978 TTCAACTATCTCTGTCCAGCATCATCAACCAATGTAGTATCTTGTGATTCAT 1037
 QY 119 -----

Db 1038 ACAAGTGAAGAGGCTGAATTCCTCAAGACCTATTCATTCAAAATAGCCTGGGTT 1097
 QY 119 -----

Db 1098 GGTGTTTATAGCCATTTCATCATCATGTTCTCTGCTGCTGGGGTTATCTTAGTG 1157
 QY 119 -----

Db 1158 CTTCTCATGAATCGGGTGTTCCTCAAAATTTCTCCTGAGTTTCCTGTGGCACTGCGCTT 1217
 QY 119 -----

Db 1218 GGGACTTGTAGTGTGATGCTTTTTCACCTTCTCCATTCATGCAAGTCAACAC 1277
 QY 119 -----

Db 1278 CATAGCTATGACATGAAGAACCAAGCATGGAATGAAGAGAGCACTTTTCAGTCAT 1337
 QY 119 -----

Db 1338 CTGTCTCTCAAAACATGAGAAAGAGTCCATTATTTGATTCACGTGAGAGGTTTACA 1397
 QY 119 -----

Db 1398 GCTTAGAGAGCCGTGATTTTATGTTTCTTGTGAACATGCTTCATATGATCAACAA 1457
 QY 119 -----

Db 1458 TTTAAAGATTAAGAAAGAAAGATCAGAAAGAACTGAAATGATGATGTGAGATT 1517
 QY 119 -----

Db 1518 AAGAGCAGTTGTCCAGATGATGATTCGACTTTCACAAATGAGAGAAAGTAGATACA 1577
 QY 119 -----

Db 1578 GATGATCGAACTGAAGGCTATTTACGAGCAGACTCAAGAGCCCTCCACTTTGATTCT 1637
 QY 119 -----

Db 1638 CAGCAGCTGACGTCTTGAGAGAGAGAGATCATGATAGTCAATGATCAGAGAA 1697
 QY 119 -----

Db 1698 GTTCAATGAATATGATCCAGAGGTGCAAGAAATGCAATTCATTTCCAGAT 1757
 QY 119 -----

Db 1758 ACACTGGCCAGTCAAGCATCTTCAACCACTATGACTATCATATTTCTCAT 1817
 QY 120 -----

Db 1818 CATCAACCAACCAAAACACCATCTTCAAGTCAAGCAGCGCTACTCTCGGAGAGAG 1877
 QY 134 LeuLysAspAlaGlyValAlaThrLeuAlaTrpMetValIleMetGlyAspGlyLeuHis 153

Db	CTGAAGAAGTCCCGGCTGCCACTCTGGCCCTGGATGCGATTAATGAGCGATGAGCCTGCAC	193
Qy	1878	
Qy	154	173
Db	1938	1997
Qy	174	193
Db	1998	2057
Qy	194	213
Db	2058	2117
Qy	214	233
Db	2118	2177
Qy	234	253
Db	2178	2237
Qy	254	273
Db	2238	2297
Qy	274	293
Db	2298	2355
Qy	294	302
Db	2358	2384
RESULT 5		
LOCUS	AX465588	3461 bp DNA linear PAT 16-JUL-2002
DEFINITION	Sequence 1 from Patent WO0216939.	
ACCESSION	AX465588	
VERSION	AX465588.1	GI:21899891
KEYWORDS		
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	1	
AUTHORS	Mack, D., Gish, K.C. and Wilson, K.E.	
JOURNAL	Method of diagnosis of cancer and screening for cancer modulators	
FEATURES	Method of diagnosis of cancer and screening for cancer modulators	
Source	Patent: WO 0216939-A 1 28-FEB-2002;	
	EOS Biotechnology, Inc. (US)	
	Location/Qualifiers	
	1..3461	
	/organism="Homo sapiens"	
	/mol_type="unassigned DNA"	
	/db_xref="taxon:9606"	
	/note="Prostate/bladder/ovarian/bladder cancer protein	
	human BCR4 CDNA"	
	138..2405	
	/note="unamed protein product; human BCR4"	
	/codon_start=1	
	/protein_id="CAD42374.1"	
	/db_xref="GI:21899892"	
	/translation="MARKUSLILITPALTSTNPLHELKAAPPTTEKISPNMESGII	
	EHSDHERSHDEHSHDHHSHHHNHASGKKRKLACDHDSDSGKOPKDR	
	SOGRKARPEHASGRNARVSASVSTVNTVSEGTLETLETPEPKLPPKDV	
	SSPTVSTKSRVSRLAGRTNESVSPRGFMYSRNTNENPOCFNASKLISHGK	
	GIQVPLNTEFVLCPLAINOIDASCLIHSEKKAELPPTYSLQIMVGGFALS	
	ISPSLIGLIVPLKMRVPEFKLSFLVALVGTSGDAPHLIPHSASHSHSHSH	
	EPAMKRPGLPSHLSQNIIESAFAFDSFTWKGLTGLGYLEMYLVENLTLIKQPKD	
	KKQOKRENDVDVETIKQLSTESQSLSTNEKVDTDRTGYLELADSQESHFPOQ	
	PAVLEESEVILAAHPQEVNVEYVGRGKKNCHSHFHDTLQSDLLHHNDYHHL	
	HHNHNHHPHSHSQRYREELKDAQVAVLAVMVGIDLHNFSGLAIGAAFTGSL	

ALIGNMENT		GLTSSVAVFCHLELHELGLDPVLLKAGTGYQAVLYNNAASMLAVLGMATGTFIIGHVY ENVSMWIFALTAGLPMYVALVDWVPEMLHNDASDHGSRGWYFFLQNNAGMLGFGIMT LISIFEKIVERINF"	
ORIGIN			
Alignment Scores:			
Pred. No.:	1,39e-111	Length:	3461
Score:	1356.50	Matches:	301
Percent Similarity:	40.00%	Conservative:	1
Best Local Similarity:	39.87%	Mismatches:	0
Query Match:	84.83%	Indels:	453
DB:	6	Gaps:	1
US-10-659-004-104 (1-302) x AX465588 (1-3461)			
QY	1	MetaLaarGlyLeuSerValIleuLeuThrPheAlaLeuSerValIThrAsnPro	20
DB	138	ATGGGAGGAAGTATATCTGTAACTTGATCTCGACCTTGCCCTCTCGTCACAATCC	197
QY	21	LeuHisGluLeuValAlaAlaPheProGlnThrThrGluValIleSerProAsnTrp	40
DB	198	CTTCATGAACATAAAGCAGCTGCTTCCCGACACACTGAGAAATATGATCCGAAATGG	257
QY	41	GIUSeRGylIleAsnValAspLeuAlaIleSerThrArgGlnTrpHisLeuGlnIleu	60
DB	258	GAATCTGGCATTAAAGTTGACTGTGGCAATTTCCACACGGCAATATCATCTACACAGCTT	317
QY	61	PheTYrArgTYrGluAsnAsnSerLeuSerValGluGlyPheArgLYleLeuGln	80
DB	318	TTCTACCCCTATGGAGAAATATATCTTGTGTCAGTGAAGGTTTCAGAAATTACTTCCA	377
QY	81	AsnIleGylIleAspLYsIleLYsArgIleHisIleHisIleAspHisAspHisIleSer	100
DB	378	AATATAGGCATATAGATTAATAAAGATTCATATACCATATGACCATGACCATCATCTCA	437
QY	101	AspHisGluHisIleSerAspHisGluArgHisSerAspHisGluHisIleSerAsp---	119
DB	438	GACCAAGCATCACTCAGACCATGAGCGTCATCTAGACCATGAGCATCTCAGACGAC	497
QY	119	-----	119
DB	498	GAGCATCACTGTGACCATGATCATCTCTCACCATAATCATGCTTCTTGTTAAAAAT	557
QY	119	-----	119
DB	558	AAGGAAAGGCTCTTTGGCCAGACCATGATCTCAAGTATGTCAGTAAGATCTTAGAAAC	617
QY	119	-----	119
DB	618	AGCCAGGGGAAAGAGAGCTCACCGACAGAACATGCGCAGTGTAGAAAGATGTCAGAGAC	677
QY	119	-----	119
DB	678	AGTGTAGTGTAGTGAAGTGAACCTCAACTGTATCAAACACTGTCTGTGAAGAACTCAC	737
QY	119	-----	119
DB	738	TTTCTAGAGCAATAGAGACTCCAGACTGTGAAAACCTTTCCCAAGATGTATAGCAGC	797
QY	119	-----	119
DB	798	TCCACTCACCCAGTGTACATCAAAAGAGCCGGGTAGCCGGCTGTGTAGAAACA	857
QY	119	-----	119
DB	858	AATGAATCTGTAGTAGGCCCCGAAAGGCTTTATGTATTCAGAAACAAATGAATAAT	917
QY	119	-----	119
DB	918	CCTCAGAGTGTTCATATGCATCAAAGTACTGACATCTCATGGCATGGGATCCAGGTT	977
QY	119	-----	119

Db 978 CCGTGAATGCAAGAGTTCAATATCTCTGTCAGCCATCATCACCAATTGATGCT 1037
 QY 119 ----- 119
 Db 1038 AGATTCTGTGATTCATACAGTGAAGAAAGGCTGAATTCCTCAAGAGCTATTCA 1097
 QY 119 ----- 119
 Db 1098 TTACAAATAGCCTGGGTTGGTGTATTATAGCCATTTCATCATGATTCCTGCTCTG 1157
 QY 119 ----- 119
 Db 1158 CTGGGGTTATATTAGTCCCTCATGATGGGGTCTTTCAAAATTCCTCTGATTC 1217
 QY 119 ----- 119
 Db 1218 CTGTGGGACCTGGCCGTTGGGACCTTGAAGTGTGATGCTTTTACACCTTCCTCACAT 1277
 QY 119 ----- 119
 Db 1278 TCTCATGCAAGTCACACCATATGATCATAGCCATGAAGAACAGCAATGGAATGAAGA 1337
 QY 119 ----- 119
 Db 1338 GGACCACTTTTCAATCATCTGCTCTCTCAAAACATGAAGAAAGTCCATTGATTC 1397
 QY 119 ----- 119
 Db 1398 ACGTGAAGGCTTAACAGCTCTAGAGGCTGTATTATGATGTTCTTGTGAACATGTC 1457
 QY 119 ----- 119
 Db 1458 CTCACATTGATCAAACAATTTAAGTAAAGAAAGAAAGATCAGAAAGCTGAAT 1517
 QY 119 ----- 119
 Db 1518 GATGATGATGAGATTAAGAAAGAGTGTCCAGATGATCTCACTTCAACAAAT 1577
 QY 119 ----- 119
 Db 1578 GAGGAAAGTATGATACAGATGATGCACTGAAGGCTATTACGAGAGACTCACAGAG 1637
 QY 119 ----- 119
 Db 1638 CCTCCCACTTGAATTCAGAGAGCTGAGTCTTGAGAAAGAGGTCATGATGCT 1697
 QY 119 ----- 119
 Db 1698 CATGTCATCCACAGAGATCTACATGATATGTAACAGAGGTCAGAAATATATGC 1757
 QY 119 ----- 119
 Db 1758 CATTCACATTCACAGATACACTCGGCACATGACATCTCATTCACACATCATGAC 1817
 QY 120 ----- 127
 Db 1818 TACCATCATATTTCTCATCATATCACACCAACCAACCATCTTACAGTACAGCCAG 1877
 QY 128 AAGTYSerAArgIuLuleuYsAaPaIaGIyVaIaIaThrLeuAlaTrpMetValIle 147
 Db 1878 CGTTACTCTCGGAGAGAGCTGAAGATGCCGGCTGCCACTTTGGCCCTGAGATGGTGA 1937
 QY 148 MetGIyAaPbGIyLeuHIsaenPheSerAaPbGIyLeuAlaIleGIyVaIaIaPheThrGIu 167
 Db 1938 ATGGGATGAGCTGACCAATTTCAAGCATGGCTGAGCAATGGTCTCTTTTACTGAA 1997
 QY 168 GIyLeuSerSerGIyLeuSerThrSerValAlaValPheCyseHIsaGIuLeuProHIsGIu 187
 Db 1998 GGCTTATCAAGAGGTTAAGTACTCTGTGTGTGTCTGTATAGATGGCTCTATGA 2057
 QY 188 LeuGIyAaSPheAlaValLeuLeuYsAlaGIyMetThrValIleGIyVaIaIaLeuYr 207
 Db 2058 TTAGGATCTTGTCTGTCTTACTTAAGAGCTGAGCATGACCTTAAGAGAGCTGTCTTAT 2117

QY 208 AenAlaLeuSerAlaMetLeuAlaTyrLeuGIyMetAlaThrGIyIlePheIleGIyHIs 227
 Db 2118 AATGCATTGTCAAGCAGCTGGCGTATCTTGAATGGCAACAGAAATTTTCATGGTCAT 2177
 QY 228 TyrAlaGIuAenValSerMetTrpIlePheAlaLeuThrAlaGIyLeuPheMetHIsVal 247
 Db 2178 TATCTGAATAATGTTTCTATGTGATATTGCACTTACTGTGGCTTATTCATCATGATGT 2237
 QY 248 AlaLeuValaPheMetValProGIuMetLeuHIsaenAaPbAlaSerAaPbHIsGIyCyseSer 267
 Db 2238 GCTGTGGTGTATAGTAACTGGAATGCTGACAAATGCTGACAAATGATGCTATGACCATGATGAGC 2297
 QY 268 ArgTrpGIyYrPhePheLeuGIuAaPbAlaGIyMetLeuLeuGIyPheGIyIleMetLeu 287
 Db 2298 CGTGGGAGTATTTCTTTTACAGAAAGCTGGAGATGCTTTGGGTTTGGATTAATGTGA 2357
 QY 288 LeuIleSerIlePheGIuHIsGIyIleValPheArgIleAaPhe 302
 Db 2358 CTATTTCCATATTGAAACATAAATCGTGTCTGTATTAATTTC 2402
 RESULT 6
 AX829136 3461 bp DNA linear PAT 12-DEC-2003
 LOCUS AX829136
 DEFINITION Sequence 29 from Patent WO2059377.
 ACCESSION AX829136
 VERSION AX829136.1 GI:39838917
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 Mack,D.H., Gish,K.C. and Afar,D.
 Methods of diagnosis of breast cancer, compositions and methods of
 screening for modulators of breast cancer
 Patent: WO 02059377-A 29 01-AUG-2002;
 JOURNAL EOS Biotechnology, Inc. (US)
 FEATURES
 source
 1. 3461
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.39e-111 Length: 3461
 Score: 1356.50 Matches: 301
 Percent Similarity: 40.00% Conservative: 1
 Best Local Similarity: 39.87% Mismatches: 0
 Query Match: 84.83% Indels: 453
 DB: 6 Gaps: 1
 US-10-659-004-104 (1-302) x AX829136 (1-3461)
 QY 1 MetAlaArgIyLeuSerValIleLeuIleLeuThrPheAlaLeuSerValIleAaPbPro 20
 Db 138 ATGGGAGAGAGTATCTGTAATCTTGAATCTTGCACCTTGGCTCTCTGTCAAAATCCC 197
 QY 21 LeuHIsGIuLeuYsAlaAlaPheProGIuThrArgIuYsIleSerProAaPb 40
 Db 198 CTTCATGAACATAAAGCAGCTGCTTCCCGACAGACACTGAGAAATTAAGTCGAATGG 257
 QY 41 GIuSerGIyIleAenValaPheLeuAlaIleSerThrArgIuYsIleLeuGIuInGIu 60
 Db 258 GAATCTGCATTAAGTTGACTTGGCAATTTCCACACGGCAATATCATCTACAAACGCT 317
 QY 61 PheTyrArgIyYrGIuAaPbAaPbLeuSerValGIuGIyPheArgIyLeuLeuGIu 80
 Db 318 TTCTACCGCTATGAGGAAATTAATCTTTGTGAGTTGAAGGTTCAAGAAATTAATTCTCA 377
 QY 81 AenIleGIyIleAaPbYsIleYsArgIleHIsIleHIsAaPbHIsAaPbHIsSer 100

AUTHORS

Mezer, P.S., Rastelli, L., Hermann, J.L., MacDougall, J.R., Zhong, H.,
Caeman, S.J., Boldog, F., Shinkels, R.A., Gorman, L., Cresta, O.R.,
Meyere, K.K., Folkerts, O., Martin, G.B., Eisen, A., Spaderna, S.K.,
Vernier, C.A., Bergh, C., Spytek, K.A., DiIppio, V.A., Zernusen, B.D.,
Peyman, J.A., Ellerman, K., Stone, D.J., Grose, W.M., Alsbrook, J.P.,
Lejoly, D.M., Rieger, D.K., Burgess, C.E. and Edinger, S.

TITLE

Proteins and nucleic acids encoding same
Patent: WO 02055705-A 25 18-JUL-2002;

JOURNAL

Curagen Corporation (US)

FEATURES

Location/Qualifiers
1..3445
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 3,166-111 Length: 3445
Score: 1352.50 Matches: 300
Percent Similarity: 40.00% Conservative: 2
Best Local Similarity: 39.74% Mismatches: 0
Query Match: 84.58% Indels: 453
DB: 6 Gaps: 1

US-10-659-004-104 (1-302) x AX511618 (1-3445)

QY 1 MetAlaArgLyLeuSerValIleLeuIleLeuThrPheAlaLeuSerValThrAsnPro 20
Db 117 ATGGAGAGAAATGTTATCTGTATCTTGATCTGACCTTGGCCCTCTCTGTCACAAATCCC 176
QY 21 LeuHISGLuLeuLyAlaAlaIlePheProGlnThrGluLyIleSerProAsnTrp 40
Db 177 CTTCATGAACTAAAGACAGCTGCTTCCCGACAGACAGAAATTAATGTCGATTGG 236
QY 41 GluSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTrpHisLeuGlnIleu 60
Db 237 GAATCTGGCATTAAGTTAGTTACCTTGGCAATTTCCACACGCGCATATCATCTACAACAGCTT 296
QY 61 PheTyArgTyGlyGluAsnAsnSerLeuSerValGluGlyPheArgLyLeuLeuGln 80
Db 297 TTCTACCGCTAAGGAAATTAATCTTGTCAAGGAGGTTGAGAAATTAATCTTCA 356
QY 81 AsnIleGlyIleAspLyIleLyAspArgIleHisIleHisAspHisAspHisSer 100
Db 357 AATATATGGCATATGATTAAGATTAAGAAATTCATATACCATGACCAACATCACTCA 416
QY 101 AspHisGluHisIleSerAspHisGluArgHisSerAspHisGluHisIleSerAsp 119
Db 417 GACCAAGAGATCACTCAAGCATGAGCGTCACTCAAGCATGAGCATCACTCAAGCAGC 476
QY 119 ----- 119
Db 477 GAGCATCACTGACCATGATCATCTCACTCACTATATATGCTGCTTGTGTAAAT 536
QY 119 ----- 119
Db 537 AAGCGAAAGCTCTTGGCCAGACCATGACTCAATATGTTCAAGTTAAATCTTAGAAC 596
QY 119 ----- 119
Db 597 AGCGAGGGAAGAGAGCTACGACCAAGAAATGATGTAAGAGATGCAAGGAC 656
QY 119 ----- 119
Db 657 AGTGTAGTCTAGTAGAGTGAAGCTCAATGTTGTACACACTGTCTGTGAAGAACTCAC 716
QY 119 ----- 119
Db 717 TTCTTAGAGACAAATAGAGACTCAAGACCTGGAATACTTTCCCAAGATGTAAAGCAGC 776
QY 119 ----- 119
Db 777 TCCACTCCACCAGATGTCATCAATCAAGAGCGGGGTGAGCGGGTGTGTAGAAAAACA 836

QY 119 ----- 119
Db 837 AATGAATCTGTAGTAGAGCCCCGAAAGAGCTTATGTATTCAGAAAACAAATGAAAT 896
QY 119 ----- 119
Db 897 CCTCAGAGTGTTCATATGATCAAGAGTACATGATCTCATGAGCATGGGATCCAGGTT 956
QY 119 ----- 119
Db 957 CCGGTGAATGCAACAGATTCAACTATCTGTCCAGCCATCATCAACAAATGATGCT 1016
QY 119 ----- 119
Db 1017 AGATCTGTGATTATCATACAGTGAAGAAAGAGCTGAATCCCTCCAAAGACTATTCA 1076
QY 119 ----- 119
Db 1077 TTACAAATAGCCTGGGTGGTGGTTTATAGCATTTCCATCATCAGTTTCTGTCTGTG 1136
QY 119 ----- 119
Db 1137 CTGGGGGTATCTTATAGTCCCTCATGAATCGGCTGTTTCAAAATTTCTCTGAGTTTC 1196
QY 119 ----- 119
Db 1197 CTGTGGCATCGGCCGTTGGACCTTGATGTGATGCTTTTACACCTTCTTCACAT 1256
QY 119 ----- 119
Db 1257 TCTCATGAAGTCAACACCATATGATGATGATGAGAACGAGCATGAAATGAAAGA 1316
QY 119 ----- 119
Db 1317 GAAACATTTTCAGTCACTGTCTTTCAAACATGAGAAAGTCCATTTGATTCC 1376
QY 119 ----- 119
Db 1377 ACGTGAAGGCTTAACAGCTCTGAGAGGCTGATTTATTCATGTTCTTGTGAACATGTC 1436
QY 119 ----- 119
Db 1437 CTCACATTGATCAACAAATTAAGATTAAGAAAGAAAGATCAGAAAGAACTGAAAT 1496
QY 119 ----- 119
Db 1497 GATGATGATGTGAGATTAGAGACAGTTGTCAAATGATCTCAACTTTCAACAAAT 1556
QY 119 ----- 119
Db 1557 GAGGAGAAATGATACAGATGATGCAACTGAAGCTATTTACGAGCAGACTCACAAGAG 1616
QY 119 ----- 119
Db 1617 CCTCCCACTTGTATCTCAGAGCGCTGAGTCTTGAAGAAGAGAGTCAATGATGACT 1676
QY 119 ----- 119
Db 1677 CATGCTCATCCAGAAAGTCTACATGAATATGTACCAAGAGGTGCAAGATAAATGC 1736
QY 119 ----- 119
Db 1737 CATTCACATTTCCAGATACACTGGCGCAGTCAGATCTCATTCACACCATCATGAC 1796
QY 120 ----- 127
Db 1797 TACCATCATATTTCTCATATCATCAACCAACCAACCAACCTTCAAGTACAGCGCAG 1856
QY 128 ArgTyrSerArgGluGluLeuLyAspAlaIleThrLeuAlaTrpMetValIle 147
Db 1857 CGCTACTCTCGGAGAGAGCTGAAGATCGCGGCTGCGCATCTGTGGCTGATGTGATTA 1916

Db	1273	GGACCACCTTTTCAGTCACTCTGCTTTCTCAAAACATAGAAAGAAAGTCTATTGATTC	1332
Qy	119	-----	119
Db	1333	ACGTGAAGGCTTAACAGCTCTAGAGGCTGTATTTTCATGTTTCTTGTTGAACATGTC	1392
Qy	119	-----	119
Db	1393	CTCACATTGATCAACAACTTTTAAGATTAAGAAAGAAAAGATTCAGAAAGAAACCTGAAAT	1452
Qy	119	-----	119
Db	1453	GATGATGATGTGGAGATTAAAGAGCAGTGTCTCAAGTATGATCTCACTTTCAACAAAT	1512
Qy	119	-----	119
Db	1513	GAGGAAAGTAGATACAGATGATGAACTGAAGGCTATTACGAGCACTCACAGAG	1572
Qy	119	-----	119
Db	1573	CCCTCCCACTTGATTCTCAGAGCCTGCAGTCTTGAAAGAAAGAGTCAATGATCT	1632
Qy	119	-----	119
Db	1633	CATGCTATCCACAGAAAGTTCATCAATGAATATGTACCAGAGGCTGCAGAAATATGC	1692
Qy	119	-----	119
Db	1693	CATTCACTTCCAGATACACTCGGCGCATCAAGCATCTCATTCACCAACATATGAC	1752
Qy	120	-----HisLeuProHisSerHisSerGln	127
Db	1753	TACCATCATATTTCTCATCATGACCAACCAACCAACCATCTCCACAGTCACAGCAG	1812
Qy	128	ArgTyrSerArgGluGluLeuLysAspAlaGlyValAlaThrLeuAlaTrpMetValIle	147
Db	1813	CCCTACTCTCGGAGAGAGCTGAAGATGCCGCGCTGCACATTTGGCTGATGGTGATTA	1877
Qy	148	MetGlyAspArgGlyLeuHisAsnProSerAspArgGlyLeuAlaIleGlyValAlaThrThrGlu	167
Db	1873	ATGGGAGATGGGCTGCACAACTTCCACGATGGCCTTAGCAATTTGGTCTCTTTACTGA	1933
Qy	168	GlyLeuSerSerGlyLeuSerThrSerValAlaValPheCysHisGluLeuProHisGlu	187
Db	1933	GGTTATACAAGTGTTTAAGTACTTCTGTGCTGTGTCTGTCAATGATTTGCCCTCATGA	1993
Qy	188	LeuGlyAspPheAlaValLeuLeuLysAlaGlyMetThrValIleGlnAlaValLeuTyr	207
Db	1993	TTAGGTGACTTTGGCTGTTCTTCACTTAAAGGCTGACATACCGTTAAGCAGGCTGTCTTAT	2053
Qy	208	AsnAlaLeuSerAlaMetLeuAlaTyrLeuGlyMetAlaThrGlyIlePheIleGlyHis	227
Db	2053	AATGCAATGTCAACCCATGCTGGCGGTATCTTGGAAATGGCAACAGAAATTTTCATGGTCA	2113
Qy	228	TyrAlaGluAsnValSerMetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisVal	247
Db	2113	TATGCTGAAAATGTTTCTATGTGATATTGGCACTTACTGCTGGCTTATTCATGATGGTT	2177
Qy	248	AlaLeuValAspMetValProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSer	267
Db	2173	GCCTCGTGTATATGTAACTGAAGCTGCAACAAATGATGATGATCAACCTGATGTAGC	2233
Qy	268	ArgTrpGlyTyrPhePheLeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeu	287
Db	2233	CGCTGGGGATTTCTTTTACAAATGCTGGGATGCTTTTGGGTTTGGAAATTATGTTA	2293
Qy	288	LeuIleSerIlePheGluHisValIleValPheArgIleAsnPhe	302
Db	2293	CTTATTTCCATATTGAAACATAAATGCTGTTCGTATAAATTTTC	2337

LOCUS	CO854111	2265 bp	DNA	linear	PAT 23-AUG-2004
DEFINITION	Sequence 13 from Patent WO2004067564.				
ACCESSION	CO854111				
VERSION	CO854111.1	GI:51510138			
KEYWORDS					
SOURCE					
ORGANISM	synthetic construct				
REFERENCE	1 Law, D., Gish, K.C., Murray, R. and Culp, P.				
AUTHORS	Compositions against cancer antigen liv-1 and uses thereof				
TITLE	Patent: WO 2004067564-A 13 12-AUG-2004;				
JOURNAL	PROTEIN DESIGN LABS, INC. (US)				
FEATURES	location/Qualifiers				
source	1..2265				
	/organism="synthetic construct"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:32630"				
	/note="BCR4 mutant cDNA"				
ORIGIN					
Alignment Scores:					
Pred. No.:	5.86e-109	Length:	2265		
Score:	1324.50	Matches:	297		
Percent Similarity:	39.47%	Conservative:	1		
Best Local Similarity:	39.34%	Mismatches:	4		
Query Match:	82.83%	Indels:	453		
DB:	6	Gaps:	1		
US-10-659-004-104 (1-302) x CO854111 (1-2265)					
QY	1 MetAlaArgLYLeuSerValIIleuIIleuThrPheAlaLeuSerValThrAnPro	20			
Db	1 ATGGCGAGGAAGTTATCTGTATCTTGATCTCGACCTTGCCCTCTGTCAAAATCCC	60			
QY	21 LeuHISgluLeuLYAlaAlaAlaPheProGlnThrTrgLYuLYglSerProAnThr	40			
Db	61 CTTCATGAACTAAAGCGAGCTGCTTCCCGACACCTGAAATAATTGTCGAATTTGG	120			
QY	41 GluSerGlyIIleAsnValAspLeuAlaIleSerThrArgGlnTyrlsLeuGlnLeu	60			
Db	121 GAATCGGAGATTAAAGTTGACTTGCGCAATTTCCACGGCAATATCATCTTACAACGTT	180			
QY	61 PheTyArgTyrgLYgluAsnAnSerLeuSerValGluGlyPheArgLYleuLeuGln	80			
Db	181 TTCTACCGGTATGGAGAAATATAATCTTGTGAGTTGAAGGGTTCAGAAATTAATTCTCAA	240			
QY	81 AsnIIleGlyIIleAspLYleLYAsnArgIIleHISIIleHISAspHISAspHISer	100			
Db	241 AATATATGGCATGATTAAGATTAAAGAAATTCATATACCATGACCAAGCATCTCACTCA	300			
QY	101 AspHISgluHISIIleSerAspHISgluTyrgHISerAspHISgluHISIIleSerAsp	119			
Db	301 GACCAAGAGCATCACTCAAGCATGAGCGTCACTGACCAATGAGCATCACTCAGACCCAC	360			
QY	119 -----	119			
Db	361 GAGCATCACTGTGACCATGATCATCTCACTCAACCATTAATCATGCTGTTGTGTAATAAT	420			
QY	119 -----	119			
Db	421 AAGCGAAAGCTCTTTGCCCGACCATGACTGATCAGATTAAGTTCAAGTTAAAGATCTTAGA	480			
QY	119 -----	119			
Db	481 AGCCAGGAGGAAGAGAGCTACCGACAGAAATGTCAGTGTAGAAAGAAATGTCAGGAC	540			
QY	119 -----	119			
Db	541 AGTGTAGTGTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	600			
QY	119 -----	119			

```

Db      601 TTCTAGAGACATAGAGACTCCAAGACCTGGAATACTTCCCAAGATGTAGACGC 660
Qy      119 -----
Db      661 TCACCTCCACCAAGTGTACATCAAAAGACCGGGTGACCGGCTGGCTGTAGAAAAA 720
Qy      119 -----
Db      721 AATGATCTGTAGAGAGACCCGAAAAAGGCTTTATGTATTCAGAAAACAATGAAAT 780
Qy      119 -----
Db      781 CCTCAGAGTGTTCATATGATCAAAAGCTACTGATCATGTGCGATGCGATCAGGTT 840
Qy      119 -----
Db      841 CCGGTGAATGCACAGAGTTCATATCTGTCTCCAGCATCATCAACCAATTGATGCT 900
Qy      119 -----
Db      901 AGATCTTGTCTGATTCATACAAAGTGAAGAAAGGCTGAATCCCTCCAAAGACTATTCA 960
Qy      119 -----
Db      961 TTACAAATAGCGCTGGGTTGGTGTATTATAGCAATTCATCATCAGTTCTGTCTGTG 1020
Qy      119 -----
Db      1021 CTGGGGGTTATCTTAGTGCCTCTCATGAAATCGGGTGTTCCTGTAAGTTTC 1080
Qy      119 -----
Db      1081 CTTGTGCACTGGCGGTTGGGACTTTGAGTGTAGTCTTTTACACCTTCTCCACAT 1140
Qy      119 -----
Db      1141 TCTCATGCAAGTCAACCATAGTCATAGCATGAGAACAGCAAGATGGAATGAAAAA 1200
Qy      119 -----
Db      1201 GGACCACTTTTCAGTCATCTGTCTTCATAAACAATAGAAAGAGCTTATTTGATTC 1260
Qy      119 -----
Db      1261 ACGTGAAGGCTTAACAGCTCTAGAGAGGCTGTATTTTCATGTTCTTTGTAACATGTC 1320
Qy      119 -----
Db      1321 CTCACATTGATCAAAACAATTAAAGATAAGAAAGAAAGATCAGAAAGAACTGAAAT 1380
Qy      119 -----
Db      1381 GATGATGATGTGAGATTAAAGACAGTGTCTCAAGTATGATCTCAACTTTCAAAAT 1440
Qy      119 -----
Db      1441 GAGAGAAAAGTAGATACAGATGATGAACTGAAAGCTATTTACAGACAGACTCAAGAG 1500
Qy      119 -----
Db      1501 CCCCTCCCACTTGATTCACAGACGCTGCAGCTTTGAAAGAAAGAGTCAATGATAGCT 1560
Qy      119 -----
Db      1561 CATCTCATCCACAGAGAGTCTACAAATGATATGACCAAGAGGTGCAAGATTAATGC 1620
Qy      119 -----
Db      1621 CATTCACTTTCCAGGATACACTGGCCAGTCAAGCATTCATTCACCAACCATCATGAC 1680
Qy      120 -----
Db      1681 TACCATATATTCTCATCATACCAACCAAAACCAACCATCTCTCAAGTCAACAGCCAG 1740

```

```

Qy      128 ArgTyrSerArgGluGluLeuValAspAlaGlyValAlaThrLeuAlaThrMetValIle 147
Db      1741 CGTACTCTCGGGAGAGAGTGAAGATGCCGGCGCTCCGCACTTTGGCGCTGATGTAGTA 1800
Qy      148 MetGlyAspGlyLeuHisAsnPheSerAspGlyLeuAlaIleGlyValAlaPheThrGlu 167
Db      1801 ATGGGTATAGGCTCGCACAATTTTCAGCGATGGCTTACCAATTTGGTCTTACTGAA 1860
Qy      168 GlyLeuSerSerGlyLeuSerThrSerValAlaValPheCysHisGluLeuProHisGlu 187
Db      1861 GGCCTTACAAAGTGTAAAGTAACTTCTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCA 1920
Qy      188 LeuGlyAspPheAlaValIleLeuValAlaGlyMetThrValIleValIleValIleThr 207
Db      1921 TTAGGTACCTTGTGTGTCTTAAAGCTGGCAGTGAACCGTTAAGCAGGCTGTCTTAT 1980
Qy      208 AsnAlaLeuSerAlaMetLeuAlaTyrLeuGlyMetAlaThrGlyIlePheIleGlyHis 227
Db      1981 AATGCAATTTGCAGCATGCTGGCTATCTTGAAATGGCAACGAAATTTTCATGTGTCAT 2040
Qy      228 TyrAlaGluAsnValSerMetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisVal 247
Db      2041 TATGCTGAAATGTTTCTATGTGATATTGCACTTACTGCTGCTTATTCATGATGTT 2100
Qy      248 AlaLeuValAspMetValProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSer 267
Db      2101 GCTCTGTGTATTAAGTAACTGGAATCTGCACATATGCTATGATGACCATGATGTAAGC 2160
Qy      268 ArgTrpGlyTyrPhePheLeuGluAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeu 287
Db      2161 CGCTGGGGGTTATTTCTTTTACAGAAATGCTGGATGCTTTGGGTTTGGAAATATGTGA 2220
Qy      288 LeuIleSerIlePheGluHisIleValIlePheArgIleAsnIle 302
Db      2221 CTTATTTCCATATTGGAACATTAATCGTGTTCGTATTAATTTTC 2265

RESULT 10
AX207205 3461.bp DNA linear PAT 30-AUG-2001
LOCUS Sequence 1 from Patent M00155178
DEFINITION AX207205
ACCESSION AX207205
VERSION AX207205.1 GI:15394960
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1. Goddard, A., Gurney, A.L., Smith, V., Hongo, J.A. and de Sauvage, F.
   Compositions and methods for treatment of cancer
   Patent: WO 0155178-A 1 02-AUG-2001;
   GENENTECH, INC. (US)
FEATURES
    source
        1..3461
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1, Size: 107 Length: 3461
Score: 1311.50 Matches: 299
Percent Similarity: 40.05%
Best Local Similarity: 39.92%
Query Match: 82.02%
DB: 6 Gaps: 1

US-10-659-004-104 (1-302) x AX207205 (1-3461)
Qy      1 MetAlaArgGlyLeuSerValIleLeuIleLeuThrPheAlaLeuSerValThrAsnPro 20
Db      138 ATGGCGAGAGAGTATCTGTATCTTATCTGATCTGACCTTGGCCCTCTGTGTACAAATCCC 197

```

QY 21 LeuH1sg1uLeuVal1a1a1aPheProGlnThrGluVal1eSerProAsnTrp 40
 Db 198 CTTGATGAACCTAAAGCAGCTGCTTCCCGACACCTGAAGAAATTAAGTCCGATTTGG 257
 QY 41 GluSerGly11eAsnVal1aPheVal1a1eSerThrArgGlnTrpHisLeuGlnLeu 60
 Db 258 GAATCTGGCATTAAGTTGAGCTTGCAATTTCCACAGCAATATCATCTCAACACGCTT 317
 QY 61 PheTrArgTrpGlyGluAsnAsnSerLeuSerVal1GluGlyPheArgLysLeuLeuGln 80
 Db 318 TTCTACCCCTATGGAGAAATTAATCTTGTGAGTTGAAGGTTCAAGAAATTAATCTTCA 377
 QY 81 Asn11eGly11eAspLys11eLysArg11eHis11eHis1aAspHisAspHisSer 100
 Db 378 AATATAGGCATATGATTAAGATTAAAGAAATCCATATACACCATGACACACCATCATCA 437
 QY 101 AspHisGluHis1eSerAspHisGluArgHisSerAspHisGluHisSerAsp 119
 Db 438 GACCAAGCAGCATCTGACAGACCATGAGCGTCACTCAGACCATGACATCACTCAGACAC 497
 QY 119 ----- 119
 Db 498 GAGCATCACTGACCATATATCATGCTGCTTGTGTAATAAATAAGCGAAAGCTTTGC 557
 QY 119 ----- 119
 Db 558 CCAGACATGACTGATAGTTCAGGTAAAGTCTTAAGAAACAGCCAGGGAGAAAGAGCT 617
 QY 119 ----- 119
 Db 618 CACCGACCAAGACATGCCAGTGTAGAAAGAAATGTCAGAGACATGTTAGTCTAGTAA 677
 QY 119 ----- 119
 Db 678 GTGACCTCACTGTGTACAAACACTGTCTGTGAAGAACTCACTTCTAGAGACATAGAG 737
 QY 119 ----- 119
 Db 738 ACTCCAGACCTGGAATACTTCTCCCAAGATGTAGAGAGCTCCACTCACCCAGTGTTC 797
 QY 119 ----- 119
 Db 798 ACATCAAGAGCCGGGTGAGCCGGCTGCTGTAGGAAAACAAAATGAATCTGTGAGTGAG 857
 QY 119 ----- 119
 Db 858 CCCCCAAAAGCTTTATGTATTCAGAAACAGAAATGAATAATCTCGAGAGTGTTCAT 917
 QY 119 ----- 119
 Db 918 GCATCAAGCTACTGACATCTCATGCGCATGCGCATCCAGTTCGCGTGAATCAACAGAG 977
 QY 119 ----- 119
 Db 978 TTCAACTATCTGTGCCAGCCATCATCAACAAATTAATGATGATCTTGTCTGATTCAT 1037
 QY 119 ----- 119
 Db 1038 ACAAGTGAAGAAAGCTGAATAATCCCTCAAGAAAGCTATTCATTAACAATAGCTGGGTT 1097
 QY 119 ----- 119
 Db 1098 GGTGTTTTATAGCCATTTCCATCATCAGTTTCTGTCTGTGCGGGGTATTTCTTAGTG 1157
 QY 119 ----- 119
 Db 1158 CCTTCATGAATCGGGTGTTCCTCAAAATTTCTCGAGTTTCTGTGTCGACTGGCCGTT 1217
 QY 119 ----- 119
 Db 1218 GGGACTTTGAGTGTGATGCTTTTTCACACTTTCTTCACATTTCTCATGCAAGTCAACAC 1277

QY 119 ----- 119
 Db 1278 CATATGATATGCCATGAAGAACCAATGAAATGAAGAGACCACTTTCAATCAT 1337
 QY 119 ----- 119
 Db 1338 CTGTCTTCTCAAAACATAGAAGAAAGTCCATTTTGTATTCACGTGGAAGGTCTAACA 1397
 QY 119 ----- 119
 Db 1398 GCTCTAGAGCCCTGTATTTTCATGTTCTTGTGTAACATGCTCCATTTGATCAACAA 1457
 QY 119 ----- 119
 Db 1458 TTTAAAGTAAAGAAAGAAAGATCAGAGAAACCTGAAATGATGATGTGAGATT 1517
 QY 119 ----- 119
 Db 1518 AAGAGCAGTTGTCCAGTATGATTCATCTCAACTTTCACAAATGAGAGAAAGTAGATCA 1577
 QY 119 ----- 119
 Db 1578 GATGATGAACTGAAGCTATTTACGACAGACTCAGAGAGCCCTCCACTTGTATTC 1637
 QY 119 ----- 119
 Db 1638 CAGAGCCTGCACTTTGGAAGAAAGAGTCAATGATGCTCATGCTCATCCACAGAA 1697
 QY 119 ----- 119
 Db 1698 GTTACATGAATATGTATCCAGAGGGGTCAAGAAATAATGCCATTCATTTCCAGAT 1757
 QY 119 ----- 119
 Db 1758 AACTCGGCAGTGCAGATCTCATTCACCATCATGACTACCATATATTCCTCAT 1817
 QY 120 ----- 133
 Db 1818 CATCAACCAACCAAAACACCATCTCAGCTCACAGCCAGCGCTACTCTCGGAGAGAG 1877
 QY 134 Leu1aAspAlaGlyVal1a1aThrLeu1a1aTrpMetVal11eMetGlyAspGlyLeuHis 153
 Db 1878 CTGAAGATGCGCGGTGCCACTTGTGCGCATTTGCGTGAATGATGAGTGGATGCGCTGCAC 1937
 QY 154 AsnPheSerAspGlyLeuAla11eGlyAla1aPheThrGluGlyLeuSerSerGlyLeu 173
 Db 1938 AATTCAGCGATGCGCTGACATTTGCTCTTTTACTGAAGGCTTATCAAGTGTATA 1997
 QY 174 SerThrSerVal1a1aVal1aPheCysHisGluLeuProHisGluLeuGlyAspPheAlaVal 193
 Db 1998 AGTACTTGTGTGCTGTGTGTCTGTCAATGAGTTGCTCATGAATTAAGTGAATTTCTGTT 2057
 QY 194 LeuLeuVal1aGlyMetThrVal1aGln1aValLeuTrpAsnAlaLeuSerAlaMet 213
 Db 2058 CTACTAAAGCGTGGATGACCGGTATAGCAGCGTGTCTTTTAATGATTTGCAACCATAG 2117
 QY 214 LeuAlaTrpLeuGlyMetAlaThrGly11ePhe11eGlyHisTrpAlaGluAsnValSer 233
 Db 2118 CTGGCGTATCTTGGATGCAACAGAAATTTTCATTTGCTCAATTAAGCTGAAGAAATGTTCT 2177
 QY 234 MetTrp11ePheAlaLeuThrAlaGlyLeuPheMetHisVal1a1aLeuVal1aAspMetVal 253
 Db 2178 ATGTGATATTTGCACTTACGTGCTGCTTATTCATGATGATGCTGTGATGATTTGTA 2237
 QY 254 ProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSerArgTrpGlyTrpPhePhe 273
 Db 2238 CCTGAATGCTGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2297
 QY 274 LeuGlnAsnAlaGlyMetLeuLeuGlyPheGly11eMetLeuLeu11eSer11ePheGlu 293
 Db 2298 TTACAGAAATGCTGGATGCTTTTGGTGTGGAATTAATGATTAATTTCAATTTTCA 2356
 QY 294 HisLys11eVal1aPheArg11eAsnPhe 302

Db	2357	CATATAAATCGTCTTCGT-ATAAATTTC	2382
RESULT 11			
LOCUS	BC055012		
DEFINITION	Mus musculus solute carrier family 39 (metal ion transporter), member 6, mRNA (CDNA clone MGC:62673 IMAGE:6373769), complete cds.		
ACCESSION	BC055012		
VERSION	BC055012.1		
KEYWORDS	GI:32822908		
SOURCE	MGC.		
ORGANISM	Mus musculus (house mouse)		
REFERENCE	Mus musculus		
AUTHORS	Eukaryota Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 3610)		
	Strauberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shemmen, C.M., Schuler, G.D., Altschuld, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshitsuyuki, S., Gracini, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Muller, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Woley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulik, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalka, U., Smalins, D.E., Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.		
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
MEDLINE	22388257		
PUBMED	12477932		
REFERENCE	2 (bases 1 to 3610)		
AUTHORS	Strauberg, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (11-JUL-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: http://mgc.ncl.nih.gov		
COMMENT	Contact: MGC help desk Email: cgapb-r@mail.nih.gov Tissue Procurement: Susan L. Sullivan, PhD. CDNA Library Preparation: Resgen, Invitrogen Corp CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisc.nih.gov/ Contact: nisc.mgc@nih.gov Ahter, N., Ayala, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakeley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, U., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R., Maduro, Q.L., Maestriello, C., Maseri, B., Mastriani, J.C., McCloskey, J.C., McConell, J., Pearson, R., Stanciroop, S., Thomas, P.J., Touchman, J.W., Tarragon, C., Vogt, J.L., Walker, M.A., Weheby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.		
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAX Plate: 116 Row: 9 Column: 18 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 20982834. 1..3610		

```

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/cclone="MGC:62673 IMAGE:6337369"
/clisue type="Olfactory epithelium, neonatal mouse,
CS7B1/6"
/clone_lib="NIH MGC_129"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"
1..3610
/gene="Slc39a6"
/note="synonym: Ermelin"
/db_xref="locusid:106957"
/db_xref="MGI:2147279"
211..2508
/gene="Slc39a6"
/codon_start=1
/product="Slc39a6 protein"
/protein_id="AAHS5012.1"
/db_xref="GI:32822909"
/db_xref="LOCUSID:106957"
/db_xref="MGI:2147279"
/translaton="MATDLSVMILTFPALWTSPLHELOSTAAFSQTEKINSMEBCG
VNDAIATWMORHLQQLFVRVGENDSLSVGPRLKIONIGDIKXIKVIHHDEHNAD
HENHSDHEHNSDHEHNSDHEHNSDHEHNSDHEHNSHTYAGKNRKACPDLDSON
SGNRPRTSLDGSGSPAEHMNGRRNIKESASSVTSAVINNVASEGTRPVETITPKPG
RRTADVNPSTPTTSEKSRVGRSLARKKSNSVSHPKSFMYSTRNTDNIOCEPCT
TKLTHSHGMSIQALNLNTEFNVCAPAINOIDARACLIHTASEKKEI PKTYSLOIA
MLGEFIATSIISFLIGLVLPDMRVFPFKLFSLFVALAVGTLSGDALHLLPSH
ASHOHSHSHEPEAMEMERGBLPSHLSAONIEESYDSPTWKGLTALGGIYFMVLVEHV
LTLIKOPDKKKKKOKKPENDDEDVESKQJLSKTDSLSNEEKYDPREPRESTLRADS
QESPSPSQOPTMLEEEEVMI AHAPHDEVNEYVPCKNCHSHFDITLGQSDDLIH
HHHDYHIIHHHHHONHHPHSHORSREELKDAGIATLAAMYIMGGGLNFSDGLAI
GAATHIELSSGLSTSVAVFCHLEPHEIGEDPAVLKGMTRYKQAVLYAALSAMAYLQMA
ATGFIFHYAENVSMWIFALTAGLFMVVALVDVPMELNHADSDHGSRGYFFLOQA
GILLGRFMILLISTFEKIYFRINF"
2011..2421
/misc_feature
/gene="Slc39a6"
/note="ZIP; Region: ZIP Zinc transporter. The ZIP family
consists of zinc transport proteins and many putative
metal transporters. The main contribution to this family
is from the Arabidopsis thaliana ZIP protein family these
proteins are responsible for zinc uptake in the plant.
Also found within this family are C. elegans proteins of
unknown function which are annotated as being similar to
human growth arrest inducible gene product, although this
protein is not found within this family"
/db_xref="CDD:pfam02535"

```

ORIGIN

Alignment Scores:			
Pred. No.:	6,97e-101	Length:	3610
Score:	1237.50	Matches:	276
Percent Similarity:	38.43%	Conservative:	18
Best Local Similarity:	36.08%	Mismatches:	8
Query Match:	77.39%	Indels:	463
DB:	10	Gaps:	2

US-10-659-004-104 (1-302) x BC055012 (1-3610)

Oy	1	MetaLaarglyleuSerVallleLeulleuthrPheAlaleuSerVallThraAsnPro	20
Dd	211	ATGGCCACAGATTATCTGAATCATGATCTTGACCCTTGCGTTAGCAAGCCCC	270
Oy	21	LeuHisgileuleuySaLa---AlaAlaPherProGlnThrTrgUlyylleserProAn	39
Dd	271	CTTCATGAACAATCAATTAACAGCGCTTCTTCAGACTAAGAAAATAATTCAAAT	330
Oy	40	TrgUlyserGlyleuAenValAapLeuAlAleSerThrArgInGlyrHiSlaugInGln	59
Dd	331	TGGAAACCTGGTGTTAATGTGACTTGACGAGTTAACCATGACGAGCACCATCTGCACAG	390
Oy	60	LeuPheTyArGlyTrgUlyGuAsnAsnSerleuSerValGlUgUlyPhearGlyleuLeu	79

Db	391	CTATTCTAACCGCTACGGAGAGAAATGATTTCTTGCTCTGTGAAGCTTCAAGAAATTCCTT	450
QY	80	GLnAaenllegiylleapLyelleyaargllehisillehislsaepHisaepHis	99
Db	451	CAGAAACATAGGCATATGATTAAGATTAAAAAGAGTCCATATTAACCATGACCATGACCAAGCATCAT	510
QY	100	Se7aepHisglunHisIsa7aepHisglunA7rHisIsa7aepHisglunHisIsa7aep	119
Db	511	GCTGACCAAGAGATCACTCGAGCCATGAGCATCACTCGACCAAGAGATCACTCGAC	570
QY	119	-----	119
Db	571	CACGAGATCACTCGAGCCAGAGCATCTCGAGCCAGAGCATTCACACCGAGT	630
QY	119	-----	119
Db	631	CACACGTTGCTGGTAAAAACAATCGAAGCCTTTTGTCCAGACTTGACTTGATTAAT	690
QY	119	-----	119
Db	691	TCAGSTAAAAATCCTTAGACTTAGTCTAGGGAAAGAGTCTCGCCAGAGAGACATGAAT	750
QY	119	-----	119
Db	751	GGTAGAGAGAACATCAAGAGAGAGTCAGACTCTAGTAAAGTCACTCGCGGTATACAAC	810
QY	119	-----	119
Db	811	GTTGTCTCTGAAGAACTCGCTTTGTAGAGCAATAGAGACTCCAAAACTGGAGAGCG	870
QY	119	-----	119
Db	871	ACCAAGATGTAAACCTTTCTATCCCAACCCAGACATCAAGAGAAAGCCAGTGGCCGG	930
QY	119	-----	119
Db	931	CTGAGTCGGCTAGTAGAAGAAAGCAATGATGCTGTGAGTGAGCCAGAAAGCTTT	990
QY	119	-----	119
Db	991	ATGTATCCAGAAACAACAATGACAAATATTACAGAGTGTTCATATACACCAAGCTCTG	1050
QY	119	-----	119
Db	1051	ACATCCATGSGATGAGATCCAGGCTCTGTTGAATGCAAGAAATTAATATCTGTC	1110
QY	119	-----	119
Db	1111	CCAGCATCATCAATCAATATGATGCTCGGCTGTCTGATTCATATACCAAGTAGAAG	1170
QY	119	-----	119
Db	1171	AAAGCAGAAATCCCTCCAAAGACTTATCTTTACAAATAGCTGGCTTGGTGTTCA	1230
QY	119	-----	119
Db	1231	GCCATTCATCATGACTTCTCTGTCTCTGTAGGATCATCTTGTCACATCAAGAC	1290
QY	119	-----	119
Db	1291	CGGATATTTTCAAGTTCTCTGTGAGCTTCTGTGGCGCTGGCCGTCGAAACGCTGAGT	1350
QY	119	-----	119
Db	1351	GGGAGTCTCTGTACATCTTCTCCACACTCATGCAAGTCATCAGACAGTATAGC	1410
QY	119	-----	119
Db	1411	CATGAAGAGCAGAGATGAAATGAAGAGGCCCTTGTACGACACCTGTGGCTGAG	1470
QY	119	-----	119

D	b		1471	AATRTAGAAAGAAGCTCATTGTTTGAATTCACAGTGAAAAGGTCTGACGGCTTAGGGGC	1530
O	y		119	-----	119
D	b		1531	TTAATATTCATGTTCTTGTGTGAACAAGTACTCATCGATCAAGCAATTTAAGATAAG	1590
O	y		119	-----	119
D	b		1591	AAAAAGAAAGATCAAAAAGAAAGCTGAAAATGATGAGAGTGTGAGAACAAAGAACGCTG	1650
O	y		119	-----	119
D	b		1651	TCCAAATAGACTCTCGACTTTCTCTCAAATGAAAGAGAGGTGACCAGGGGAACGACT	1710
O	y		119	-----	119
D	b		1711	GAAAGTATCTGCGAGCCGACTCCCAAGACCCTGCCCTTTGATTTCCAGACGCCGACG	1770
O	y		119	-----	119
D	b		1771	ATGTTGGAAGAAGAGAGGTCAATGATAGCCCATGACACCCACAAGAGCTTCATGAA	1830
O	y		119	-----	119
D	b		1831	TATGTGCCCAAGGGGCTGCAGAAACAAGTGCATTCACTTCACAGATACGTGGCCAG	1890
O	y		119	-----	119
D	b		1891	TCCGACGACCTCATCCACCAACATCACAGACTACATCATCTTGACACCAACCAAC	1950
O	y		120	-----HSHSPROHSERHISERGLARGLTYRSERGLUGLLEULYASPAALA	137
D	b		1951	CAGAACCAACCCCTCCACAGCCACAGGCGAGGCTACTCTGAGAGAGCTGMAAGAGCC	2010
O	y		138	GIYVALAAlATHrLeuAlATrPMeValIImeRGlyAspGlyLeuHiAsenPeSerAsp	157
D	b		2011	GGCATGTGCACATTGGCTGGATGTGATCATGGGCGACGGGCTGCACAAATTCAGTGAC	2070
O	y		158	GIYLeuAlaIleGIyAlaAlaAPheThrGlUGlyLeuSeSerGIYLeuSerThrSeVal	177
D	b		2071	GGCCTTGCTATGGTGGTGCCTTCACCGAGGGTTTTGCCAGTGGCTTAAACCACTCTGTG	2130
O	y		178	AlaValAPheCyVhiIGLIleuProHISGILeuGIYAspPheAlaValLeuLYPaLa	197
D	b		2131	GCTGTGTTCTGTCAATGAAGCTGCTCATGAACATAGTACTTTGCTGTTTGTCTAAAGCT	2190
O	y		198	GIYMetThrValLySGInALaValLeuTYRAsnAlaleuSerAlameLleuAlaTYRLeu	217
D	b		2191	GGCATGACTGTCAACAGCAGCTGTCTCTAATATCCTGTGCAGCAGCATGTTGGCTTACT	2250
O	y		218	GIYMetalathrgIylEPheIIeGIYHIstYRlaGluAnValIserMetTrIPhe	237
D	b		2251	GGAAATGGCAACAGGGATATTCATCGGGCATTTATCAGAAAATGTTTCTATGTGATATTC	2310
O	y		228	AlaLeuThrAlaGIYLeuPheMetHIVAlaAlaLeuValAPheMetValProGluMetLeu	257
D	b		2311	GCACCTACCTGCGCGGTGTGTCATGATATGCTCTGTGTGACATGAGTGCCTGAAGTTG	2370
O	y		258	HIsenPeRALasErAspHiEGlyCyseSerArgTrpGIYTYRPhePheLeuGINAsnAla	277
D	b		2371	CACAAATGATGCTAAGTACACAGATCAGCGCGTTGGGGATATTTCTTCTCAGAAATGCT	2430
O	y		278	GIYMetLeuLeuGIYPhegIYLewetLeuLeuIleserIIepheGluHisLYelleyAl	297
D	b		2431	GGGATACATTCCTCGGTTTTGGAATATGTTACTCATTTCCATATTTGAGCATMAAATGTG	2490
O	y		298	PHeATgIIeAsnPhe302	
D	b		2491	TTTCGTATAAATTTC2505	
<hr/>					
RESULT 12					
AB071697					
LOCUS					
AB071697 3287 bp mRNA linear ROD 15-MAR-2002					

DEFINITION Mus musculus ermelin mRNA for endoplasmic reticulum membrane protein, complete cds.

ACCESSION AB071697

VERSION AB071697.1 GI:19570345

KEYWORDS Mus musculus (house mouse)

SOURCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

ORGANISM Mus musculus

REFERENCE 1 Suzuki, A. and Endo, T.

AUTHORS Ermelin, an endoplasmic reticulum transmembrane protein, contains the novel HELP domain conserved in eukaryotes

TITLE Gene 284 (1-2), 31-40 (2002)

JOURNAL 21886618

MEDLINE 11891044

PUBMED 11891044

REFERENCE 2 (bases 1 to 3287)

AUTHORS Endo, T.

TITLE Direct Submission

AUTHORS Submitted (17-SEP-2001) Takeshi Endo, Chiba University, Department of Biology, Faculty of Science, 1-33 Yayoi-cho, Inage-ku, Chiba, Chiba 263-8522, Japan (E-mail: tendo@cuphd.nd.chiba-u.ac.jp, Tel:81-43-290-3911, Fax:81-43-290-3911)

JOURNAL Location/Qualifiers

FEATURES

source

1..3287

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/cell_line="C2"

/cell_type="skeletal muscle myoblasts"

/clone_id="lambda ZAPII C2 myoblast cDNA"

1..3287

/gene="ermelin"

704..2221

/gene="ermelin"

/codon_start=1

/product="endoplasmic reticulum membrane protein"

/protein_id="BAB6300.1"

/db_xref="GI:19570346"

/translation="MYSRNTNDNIQECFNTKLTSHQMSIQALLNATEFNYLCPAII
NQIDARACLIHTASEKKAIEIPKTYSLQIAMLGFIATISIFLSLGLVILPLMNRV
FFKFLSELVALAGTSLGSDALHMLPHSHSHSHSHSEEPAMKRGPLFSLHSAQ
NIESSYFDSITWKGILTALGGLFPMFLVHVHVLTKOPDKKKKKKKKRENDVEDSKK
QLSKYDSQLSNEKEVDEGRPESTLRADSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
VYNEVPERGCKNCKSHPHDTLQSDSDILHHHHDYHHLHHHHQNHHPHSHSQRYSR
BELKDAGIATLAWVIMGDGLNFSDELATIAAFTGEGSSGLSTSVAFCHDELHELG
DPAVLKAGMTVKQAVLYNALSAMLAYGMATGIFIGYANVSMWIFALTAGLPMYV
ALVDMVPEMLHNDASDHGSRMGYFFLGNIILGFGIMLISIFEHKIVRINF"

ORIGIN

Alignment Scores:

Pred. No.: 6,496-93 Length: 3287

Score: 1148.00 Matches: 260

Percent Similarity: 37.35% Conservative: 16

Best Local Similarity: 35.18% Mismatches: 6

Query Match: 71.79% Indels: 458

DB: 10 Gaps: 2

US-10-659-004-104 (1-302) x AB071697 (1-3287)

QY 21 LeuHAGLUleuLYaLa---AlaAlaPheProGInThrThrGluLYaLeSerProAsn 39

DB 3 CTTCATGAACATCAATCAACAGCTGCTTCTCTCAGACTACTGAGAAATAATCAAT 62

QY 40 TTPGUGerGlyLLeaSnValaSPLeuLaLeSerThraGInTyPHeLLeuGInGln 59

DB 63 TGGGAACCTGGTCTTAATCTTGACTTGGCAGTACACGACGACACATCTGACGAG 122

QY 60 LeuPheTyArgTyGlyGluSnAaSnSerLeuSerValGluGlyPheArgLYaLeu 79

DB 123 CTATCTACCCCTACGAGAGATGATCTCTGTCTGTGAAGGCTTCAGAAATGCTT 182

QY 80 GlnaenleGlyLLeaSPLeuLYaLeLeuHleHleHleHleHleHleHleHleHle 99

DB 183 CAGAACTAGGCAATGATTAAGATTTAAAGATCCATATACACCATGACAGACATCAT 242

QY 100 SerAaphiSGLuHleHleSerAaphiSGLuArgHleSerAaphiSGLuHleHleSerA 119

DB 243 GCTGACCAAGACATCACTCGGACCATATGATCATCTCGGACCAAGACATCACTCGGA 301

QY 119 ----- 119

DB 302 CCAGAGCATCATCTGGACCAAGACCATATTCCACCGAGTCACACGTTGCTGTAA 361

QY 119 ----- 119

DB 362 AAACATGGAAGCCTTTTGTCCAGACTTGACTTGATTAATTCAGTAAATCTTAG 421

QY 119 ----- 119

DB 422 AACTAGTGAAGAAAGATCTCGCCAGACATGATGATGAGAGACATCA 481

QY 119 ----- 119

DB 482 GAGAGTGAAGCTCTAGTGAAGTGAACCTCGCGGTATACAACGCTGTCTGAAGAAC 541

QY 119 ----- 119

DB 542 TCGCTTGTAGAGCAATAGACTCCAAACCTGGAGAGACGACCAAGATGTAACT 601

QY 119 ----- 119

DB 602 TCTACCCACCCAGCATCAGGAGAAAGCCGAGTGGCCGCGTGAATCGGTAGTAGG 661

QY 119 ----- 119

DB 662 AAGAAAGCAATGAGTCTGTAGTGAAGCCAGAAAGCTTATGTATTCAGAAACA 721

QY 119 ----- 119

DB 722 AATGACATATTCAGAGAGTGTTCATATCAACCAAGCTGTGACATCCCATGAGTAGC 781

QY 119 ----- 119

DB 782 ATCCAGGCTGTGTAATGCAACGAAATTAATATCTCTGCCAGCATCATCAATCA 841

QY 119 ----- 119

DB 842 ATTGATGCTCGGCTGTGTCTGATTCATACAGCAAGTAGAAGAAAGCAGAAATCCCTCA 901

QY 119 ----- 119

DB 902 AAGACTATTCCTTAACAATAGCCTGCTGTGCTTCATAGCCATTCATCATCACTAGT 961

QY 119 ----- 119

DB 962 TTCTGTCTCTGCTGAGAGTATCTGTGTGCACTCATGAACGGGTATTTTCAAGTTC 1021

QY 119 ----- 119

DB 1022 CTGCTAGCTTCTCTGTGCGCTGCGCTGCGAAGCTGAGTGGCGATCTCTGTACAT 1081

QY 119 ----- 119

DB 1082 CTCTCCCACTCTCATGCAAGTCATGACAGCATGATCATTAGCAAGACCGAGATG 1141

QY 119 ----- 119

DB 1142 GAATGAAGAAAGGCCCTTGTGACGCACTGTGCGTCAATATAGAAAGAAAGCTCC 1201

QY 119 ----- 119

DB 1202 TATTTGATTCACGTGAAGAGTCTGACGCTCTAGGGGCTTATATTCATGTTTCTT 1261

QY 119 ----- 119

```

Db      1262 GTGGAACAGTACTACACTGATCAAGCAATTTAAAGTAAGAAAAAGAGATCAAAAG 1321
Qy      119 ----- 119
Db      1222 AAACGTGAAATGATGAGATGTGGAGAGCAAGAACAGCTGTCCAAATAGCACTCTGAG 1381
Qy      119 ----- 119
Db      1382 CTTTCTTCMAATGAGAGAGGTGAGCCAGGGAGAACGACTGAAGCTATCTGGAGCC 1441
Qy      119 ----- 119
Db      1442 GACTCCCAAGAGCCCTCCCTTTGATTCCTCAGCAGCCGAGATGTTGMAAGAGAGAG 1501
Qy      119 ----- 119
Db      1502 GTCATGATAGCCCATGACACCCACAGAGAGTCTACATGAAATATGTGCCAGGGGCTGC 1561
Qy      119 ----- 119
Db      1562 AAGAACAGTGCATTCACATTCGACGATAGCTGGCCAGTCCGAGCACTCATCCAC 1621
Qy      120 ----- 123
Db      1622 CACCATCAGACTACCATCATCTTCTGCACACACACACACACACACACACACCTCAG 1681
Qy      124 SerHisSerGlnArgTyrSerArgGlnGlnLeuValAspAlaGlyValAlaThrLeuAla 143
Db      1682 AGCCACACAGCAGCGCTACTCTCGAAGAGAGCGCGGATTCGCCACATTCGAGCC 1741
Qy      144 TTPMetValIleMetGlyAspGlyLeuHisAsnProSerSerArgGlyLeuAlaIleGlyAla 163
Db      1742 TGGAGTGGATCATGATGAGGCGGCGCTGCACAAATTCAGTACGCGCTTGCTATTGCTGCT 1801
Qy      164 AlaPheThrGlnGlnGlyLeuSerSerGlyLeuSerThrSerValAlaValPheCysHisGln 183
Db      1802 GCCTTCACCGAGGGTGTTCCTGCGAGCGGCTAACCACTCTGTGCTGTCTGTATGAA 1861
Qy      184 LeuProHisGlnLeuGlyAspPheAlaValLeuLeuValAlaGlyMetThrValIleGln 203
Db      1862 CTGCCTCATGAAGTACTAGTGAATCTTCTGCTGTTGCTAAAGGCTGGCATGACTGTCAAGCAG 1921
Qy      204 AlaValIleTyrAsnAlaIleuSerAlaMetLeuAlaTyrLeuGlnMetAlaThrGlyIle 223
Db      1922 GCTGAGCTCTTAAGTCTCTGTGACGCAATGTGGCCCTACCTTGGAATGGCAACAGAGATA 1981
Qy      224 PheIleGlyHisTyrAlaGlnAsnValSerMetTrpIlePheAlaLeuThrAlaGlyLeu 243
Db      1982 TTCAATCGGCATTAATGCAAAAATGTTTCTATGTGATTTGCACTCCTCCGGCTTG 2041
Qy      244 PheMetHisValAlaIleuValAspMetValProGluMetLeuHisAsnAspAlaSerAsp 263
Db      2042 TTCATGATATGCTGCTGCTGTGACATGAGGCTGAGATGTTGCACATGATCTGATGAT 2101
Qy      264 HisGlyCysSerArgTrpGlyTyrPhePheLeuGlnAsnAlaGlyMetLeuLeuGlyPhe 283
Db      2102 CACGAGATGAGCCGTTGGGAGATATTTCTTCGCAAGAAATGCGGGATTAATCTTCGGTTT 2161
Qy      284 GlyIleMetLeuLeuIleSerIlePheGlnHisIleValIleValPheArgIleAsnPhe 302
Db      2162 GGAATTAATTTACTCATTTCCATATTGAGCATTAATAATGTGTTTCGTATTAATTTTC 2218

```

```

RESULT 13
LOCUS      BC008317          2212 bp      mRNA      linear      PRI 04-NOV-2003
DEFINITION Homo sapiens soluble carrier family 39 (zinc transporter), member 6,
ACCESSION  BC008317
VERSION     BC008317.2   GI:34782897
KEYWORDS   Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

REFERENCE
AUTHORS
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 2212)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Datchenko,L., Musina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Cassavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Udén,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raja,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Huylk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Faney,J., Helton,E., Kettelman,W., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blackless,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywicki,M.I., Skalska,U., Smalius,D.E.,
Scherer,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2238257
12477932
2 (bases 1 to 2212)
Strausberg,R.
Direct Submission
Submitted (25-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
On Sep 16, 2003, this sequence version replaced gi.14249878.
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chin, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywicki, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeary, Steven
Neas, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Seedei, Jacqueline
Schein, Duane Smalius, Michael Smith, Loraine Spence, Jeff Stott,
Michael Thorne, Miranda Teal, Nataesja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IPAL Plate: 22 Row: K Column: 4.
FEATURES
source
1..2212
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3343159"
/tissue_type="Eye, retinoblastoma"
/clone_id="NIH_MGC_16"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
<1..2212
/gene="SLC39A6"
/note="synonym: LIV-1"
/db_xref="locusid:25800"
<1..1127
/gene="SLC39A6"
/codon_start=3
/product="SLC39A6 protein"
/protein_id="AAH08317.2"

```

gene
CDS

		/db_xref="GI:34782898"
Df	xref="locusid:25800"	
/translation="HASNHRSHSHEPAMEKRGFLFSLSSONITEESAYFPSTWK TALGLGFMEFLVHLTLIKQFKDKKKVKOKPEVDNDVEIKKQSKEYSOJSTNEE VDTDRTEGYELNADSQEPHSFDQOPAVLEEBEWIAAHDPGEVNEYVRCKNC SFHDITGCSDDLHHNHDIYHIILHNHNQHNPSPHSQRYSRELBKDAGVTAMW IMGDILHNPSDGLAICAAFTEGLSGLSVVFCHLPHEDGDFVLKACGTVMQ LVYNLASMLAVIGMAATGFIGHYVENSMFIPLTAGIFMYVALVDMPEMLHND DHGCRRNGCYFFIQNNGMLLRFGIMILLISIFEKIYPRINF"		
	63..371	
misc_feature	/gene="SLC39A6"	
/note="ZIP; Region: ZIP zinc transporter. The ZIP family consists of zinc transport proteins and many putative metal transporters. The main contribution to this family is from the Arabidopsis thaliana ZIP protein family these proteins are responsible for zinc uptake in the plant. Also found within this family are C. elegans proteins of unknown function which are annotated as being similar to human growth arrest inducible gene product, although this protein is not found within this family"		
	/db_xref="COD:pfam02535"	
	594..1013	
misc_feature	/gene="SLC39A6"	
/note="ZIP; Region: ZIP zinc transporter. The ZIP family consists of zinc transport proteins and many putative metal transporters. The main contribution to this family is from the Arabidopsis thaliana ZIP protein family these proteins are responsible for zinc uptake in the plant. Also found within this family are C. elegans proteins of unknown function which are annotated as being similar to human growth arrest inducible gene product, although this protein is not found within this family"		
	/db_xref="COD:pfam02535"	
ORIGIN		
Alignment Scores:		
Pred. No.:	1,69e-80	Length: 2212
Score:	1007.00	Matches: 196
Percent Similarity:	91.32%	Conservative: 4
Best Local Similarity:	89.50%	Mismatches: 17
Query Match:	62.98%	Indels: 3
Df:	9	Gaps: 1
US-10-659-004-104 (1-302) x BC008317 (1-2212)		
Oy	86 Lysilelyvavxglienhielenhihsaerphihiseraerhis-----gu 103	
Df	469 AGAATAAATGCCATTTCACAATTTC-CACGAATACCTCGCCAGCAACAGCATTCATTCAC 527	
Oy	104 HishiserAerphiGLIvarghisserAerphiGLIHishiseraphishisp-cohis 123	
Df	528 CACCATTACTACATCATTAATTTCTCCATTCATCACACCAACAAAACACATCTTCAC 587	
Oy	124 SerhisserGINarxyt-serargigluieleulyaarpaiaagiValaaThrleuala 143	
Df	588 AGTCACACCCAGCGCTACTCTCGGAGAAGCATGAAGAAGCGCGCGCTGCCACTTGCGCC 647	
Oy	144 Trmpevaliiemecgiyaagolyuleunhiisaanpheaserapolyueuaaiiegiyla 163	
Df	648 TGGATGGGATATAGGTGATGGCTGCACAAATTCAGCGAATGGCTTACCAATGGTGGT 707	
Oy	164 AlaphethrglugilyeusersergilyeuSerThrservalalavalPhocyshiglu 183	
Df	708 GCTTTACTGAAGGCTTANCAAGTGGTTAAAGTACTTCCTGTGGCTGTGTCTGTACAG 767	
Oy	184 Leup-rohiegilyuarpheraalavalileuleulyalaigiymetThrvallyvgin 203	
Df	768 TTGCTCTCAATTAAGTGACTTTGTCTTCTTAACAAAGCGTGGCATGACCGTTAAGCAG 827	
Oy	204 AAlavalleuyrAsnalaleuSerAlamelleualaytleugiywerAlathrglylie 223	
Df	828 GCCTGCTCTTATATGCAATTCGACGCACTGCGGGCATCTTGGAATGGCAACAGGAATT 887	

`/db_xref="GI:34782898"`
`/translation="HAHHHHSHSEEPAMKRGKGLPSHLSQNTIESAPYSTKMG
TALGGLFMFLVEVLTLLKQPKKKKKKKKKENDDDVEIKQLSKTSQSSTNEE
VTDIDRREGVLRASQSPHSFDDQPVVLEEEVMAAHAPQPVNVEYVRGGKRC
SHFDDTIGQSGDDTLHHNDDVHLLHHHHQNHSHSHSQSRSEELKDYVRLATLKK
IMGGHLPNSGGLAIGAAEPFHHHSSGDSVAFAFCHLPHLEDDPVLVLLKAGTVKQ
VLVYAAFLAMLAVALMGATGIFGHXAEVSNWMIETALDAGLEPVVALVDVPEMLHND
DHGSRMGVFPFLQAGMLDGLGIMLISIEHKIVRRIN"`
63. 371
`/gene="SLC39A6"`
`/note="ZIP; Region: ZIP Zinc transporter. The ZIP family
consists of zinc transport proteins and many putative
metal transporters. The main contribution to this family
is from the Arabidopsis thaliana ZIP protein family these
proteins are responsible for zinc uptake in the plant.
Also found within this family are C. elegans proteins of
unknown function which are annotated as being similar to
human growth arrest inducible gene product, although this
protein in not found within this family"`
`/db_xref="CCD:pfam0535"`
594. 1013
`/gene="SLC39A6"`
`/note="ZIP; Region: ZIP Zinc transporter. The ZIP family
consists of zinc transport proteins and many putative
metal transporters. The main contribution to this family
is from the Arabidopsis thaliana ZIP protein family these
proteins are responsible for zinc uptake in the plant.
Also found within this family are C. elegans proteins of
unknown function which are annotated as being similar to
human growth arrest inducible gene product, although this
protein in not found within this family"`
`/db_xref="CCD:pfam0535"`

Qy	Db	Qy	Db	Qy	Db	Qy	Db	RESULT 14			
224	PhellegiYH1sYrYAlaGluAsnValserMetPrlIephEAlaLeuThrAlaGlyLeu	244	PheMetHisValAlaLeuValAspMetValProGluMetLeuHisAsnAspAlaSerAsp	264	HisGlyCysSerArgTrpGlyTrpPhePheLeuGlnAsnAlaGlyMetLeuLeuGlyPhe	1008	CATGAGTATACCGCTGGGGGATTTCTTTTAAcGAAGTCGGAGATGCTTTGGGTTTT	1068	GGAATTAGTTACTTATTTCcATATTTBAAcATAAATCGTCTTTGGTATAATTTC		
888	TTcATTGGcATTATGCTGGAAATATTTCTATGcGATTTTGcACTTACTGCTGGCTTA	948	TTcATGATGATGTGCTGCTGGTGAATATGTAcCTGGAAATGCTcACAcATATGATCTAGTGAc	284	GlyIleMetLeuLeuIleSerIlePheGlnHisIleValIlePheArgIleAsnIle	1068	GGAATTAGTTACTTATTTCcATATTTBAAcATAAATCGTCTTTGGTATAATTTC	1068	GGAATTAGTTACTTATTTCcATATTTBAAcATAAATCGTCTTTGGTATAATTTC	1124	
Qy	244	Qy	244	Qy	264	Qy	1008	LOCUS			
Db	888	Db	948	Db	284	Db	1008	Q0834560			
Qy	244	Qy	244	Qy	264	Qy	1008	Q0834560			
Db	888	Db	948	Db	284	Db	1008	Q0834560			
Qy	244	Qy	244	Qy	264	Qy	1008	Q0834560			
Db	888	Db	948	Db	284	Db	1008	Q0834560			
Qy	244	Qy	244	Qy	264	Qy	1008	Q0834560			
Db	888	Db	948	Db	284	Db	1008	Q0834560			
Qy	244	Qy	244	Qy	264	Qy	1008	Q0834560			
Db	888	Db	948	Db	284	Db	1008	Q0834560			
Qy	244	Qy	244	Qy	264	Qy	1008	Q0834560			
Db	888	Db	948	Db	284	Db	1008	Q0834560			
Qy	244	Qy	244	Qy	264	Qy	1008	Q0834560			
Db	888	Db	948	Db	284	Db	1008	Q0834560			
Qy	244	Qy	244	Qy	264	Qy	1008	Q0834560			
Db	888	Db	948	Db	284	Db	1008	Q0834560			
Qy	244	Qy	244	Qy	264	Qy	1008	Q0834560			
Db	888	Db	948	Db	284	Db	1008	Q0834560			
Qy	244	Qy	244	Qy	264	Qy	1008	Q0834560			
Db	888	Db	948	Db	284	Db	1008	Q0834560			
Qy	244	Qy	244	Qy	264	Qy	1008	Q0834560			
Db	888	Db	948	Db	284	Db	1008	Q0834560			
Qy	244	Qy	244	Qy	264	Qy	1008	Q0834560			
Db	888	Db	948	Db	284	Db	1008	Q0834560			
Qy	244	Qy	244	Qy	264	Qy	1008	Q0834560			
Db	888	Db	948	Db	284	Db	1008	Q0834560			
Qy	244	Qy	244	Qy	264	Qy	1008	Q0834560			
Db	888	Db	948	Db	284	Db	1008	Q0834560			
Qy	244	Qy	244	Qy	264	Qy	1008	Q0834560			
Db	888	Db	948	Db	284	Db	1008	Q0834560			
Qy	244	Qy	244	Qy	264	Qy	1008	Q0834560			
Db	888	Db	948	Db	284	Db	1008	Q0834560			
Qy	244	Qy	244	Qy	264	Qy	1008	Q0834560			
Db	888	Db	948	Db	284	Db	1008	Q0834560			
Qy	244	Qy	244	Qy	264	Qy	1008	Q0834560			
Db	888	Db	948	Db	284	Db	1008	Q0834560			
Qy	244	Qy	244	Qy	264	Qy	1008	Q0834560			
Db	888	Db	948	Db	284	Db	1008	Q0834560			
Qy	244	Qy	244	Qy	264	Qy	1008	Q0834560			
Db	888	Db	948	Db	284	Db	1008	Q0834560			

[illegible]

OY		144	TTPMVal11IemcGlyAerpGlyLeuH1AsnpPheSerPerpGlyLeuAla1leGLyALA	163
Db		830	TGGATGTGTAAAGGGGTGATGAGCGCTGCACATTTCAGCCAGTGGCCTTAGCATTTGGCT	889
OY		164	AlaPheThrGlnGlyLeuSerSerGlyLeuSerThiSerValAlaValPheCySHISglu	183
Db		890	GCTTTACTGAAGCCTTATCAAGTGTTTAAGTACTTCTGTGGCTGTGTTGTGCATGAG	949
OY		184	LeuProHisgileuGlyAerpPheAlaValLeuLeuYsa1aglyMetThrValLysGln	203
Db		950	TTGCCCTCATGAATTAAGTGACCTTTGGCTGTTCTACTTAAGGCTGCATGACCCTTAAGCAG	1009
OY		204	AlaValLeuTyraAna1aLeuSerAlaMetLeuAlaTyrlEugLyMeFalathrGlyLe	223
Db		1010	GCTGTCTTAAATAAGCATGTTCAGCCACTGCTGGCGTATCTTGGAAATGGACAACGAAATT	1069
OY		224	PheIIeGlyHisTyra1aGluAsnValserMetTriplePheAlaLeuThrAlaGlyLeu	243
Db		1070	TTCAATGGTCATTATGCTGAAAATGTTCTATGTGATATTGGACCTTACTAGCTGGCTTAA	1129
OY		244	PheMetHisValAlaLeuValAspMetValProGluMetLeuHisAsnaAsp1aSerSerp	263
Db		1130	TTCAATGTATGTGGCTCTGTGTGATGTGTACCTGAAAGCTGCACAATAAGCTAGTAC	1189
OY		264	HISgLYCySerARXTTRIGLYTYrPhePheLeuGlnAnaAlaGlyMeLeuLeuGlyPhe	283
Db		1190	CATGATGTAGCCGCTGGGGGATTTCTTTTACAGAATGCTGGGATGCTTTTGGGTTTT	1245
OY		284	GlyIleMetLeuLeuIleSerIlePheGlnHisTyrlEva1PheArg1LeasnPhe	302
Db		1250	GGAATTATGTTACTTATTCCATTATTTGAAACATAAATCGTGTTCGTATAAATTTTC	1306
RESULT 15				
BC054780		2815 bp	mRNA	linear
LOCUS		Mus musculus endoplasmic reticulum membrane protein, mRNA (CDNA		
DEFINITION		clone IMAGE:681386), partial cds.		
ACCESSION		BC054780		
VERSION		BC054780.1 GI:32493411		
KEYWORDS				
SOURCE		Mus musculus (house mouse)		
ORGANISM		Mus musculus		
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 2815)		
AUTHORS		Straussberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuetz,G.D., Altshuler,S.F., Zeeberg,B., Buetow,K.H., Scheffer,C.F., Bac,N.K., Hopkins,R.P., Jordan,H., Moore,T., Max,S.I., Wang,J., Halello,F., Ditachenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Uddin,T.B., Toshiyuki,S., Carinci,P., Prange,C., Kahn,S.S., Loggellano,N.A., Peters,G.J., Abrimson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.D., McKenna,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Morley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Huilyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Rahcy,U., Helton,E., Keteeman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shenverson,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schultz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalski,U., Smallu,S., Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.		
TITLE		Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
JOURNAL		PROC. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
MEDLINE		22388257		
PUBMED		12477932		
REFERENCE		2 (bases 1 to 2815)		
AUTHORS		Straussberg,R.		
TITLE		Direct Submission		
JOURNAL		Submitted (01-JUL-2003) National Institutes of Health, Mammalian		

```
REMARK
COMMENT
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-rcmail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLU)
DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant.
Web site: http://genome.uiowa.edu/tom-casavant@uiowa.edu
Contact: bent-soares@uiowa.edu; tom-casavant@uiowa.edu
Bonaldo, M.F., Akabogu, I., Bair, T., Bair, J., Crouch, K., Davis, A., Fishler, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K., Scheetz, T., Smith, C., Snir, E., Tack, D., Trout, K., Walters, J., Casavant, T., Soares, M.B.
```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLU at: http://image.lnl.uow Series: Plate Row Column: 0
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 20982834.
location/Qualifiers

1 . 2815
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6813896"
/tissue_type="Brain, mouse, 13.5, 14.5, 16.5, 17.5 dpc"
/clone_id="NII BMAP_FW0"
/lab_host="DH10B"
/note="Vector: pYX-ASC"

<1 . 2815
/gene="Erme1in-pending"
/db_xref="locusID:106957"
/db_xref="MGI:2147279"

<1 . 1751
/gene="Erme1in-pending"
/codon_start=3
/product="Erme1in-pending protein"
/protein_id="AAHS4780.1"
/db_xref="GI:32493412"
/db_xref="locusID:106957"
/db_xref="MGI:2147279"
/translation="NIKSSASSSEVTSAVNAVEGRTFETPKPRRTRKVNPE
TPPSTERSRGRLSLARKSSESVSPPRSFPYSRRTNNDIOCRPTTLTHSHGG
SIQALNTGFNYLCPIAINIQIDAACLIHTASEKKATIPKTSLSQIAMLGFTAIS
ISIFSLAGIVLVPLMNRFEPFLPSVLAVAGTLSDGLHLHPHSHASHSSHSS
BEPMEMKRGPFLFSILSAONIEESSYPSTWKDGLTGLEYFMELIVEHVLTLLIKQFK
KKKKQKQRPNDEVDIESKKSKDSOLSNBEKVDPGRSPSYLRADSOEPSPDSS
OPTMLREEEVMITAHNPQEVINEYVRCKRKCHSHPHDTLGSODSLITHHHDDYHIIG
HHNNQNHRPHSHSQRYRELKDNGIATLMAMWTMGDLNFSGDLIGAFAPEGLE
SGLSNVVAFCHELPLEHGDAVALYLKAQMTKOAYLYALSAMLVLDMAGIITFGHN
AENVSMKITALTAGLFMYVALVDMVPEMLIHNDASDHCSRWGYFFLDAGILLRGIN
LLISFEKIYIRINE"

ORIGIN

Alignment Scores:

Pred. No.:		Length:	
Score:	6.55e-80		2815
Percent Similarity:	1002.00	Matches:	191
Best Local Similarity:	92.89%	Conservative:	5
Query Match:	90.52%	Mismatches:	5
	62.66%	Indels:	10
		Gaps:	2

DB: 10

US-10-659-004-104 (1-302) x BC054780 (1-2815)

Dy 92 11ehniaAsphriaAsphihihtiseraPhiegluhlieSeraPhiigludArgnis 111

Dd 1146 ATCCACCACCATCACCACTAACAT-----CACATTTCGCCAC 118

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 19, 2005, 20:53:38 ; Search time 617 Seconds
(without alignments)
2897.507 Million cell updates/sec

Title: US-10-659-004-104

Perfect score: 1599
Sequence: 1 MARKUSVILITLTFALSVTNP.....FGIMLLISIFHKIVRINF 302

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.epoc1/US10659004/runat.14032005.181234.15731/app.query.fasta_1.455
-DB=N.Geneseq.16Dec04 -QFMT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blowum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pro -NOR=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10659004.@CGN.1.1.470.@runat.14032005.181234.15731 -NCPU=6 -ICPU=3
-NO MMAP -LARGESUBSTR -NEG_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOC
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N.Geneseq.16Dec04:*
1: geneseqn1980s:*
2: geneseqn190s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1599	100.0	1101	12	ADL57158 Human NOV
2	1359.5	85.0	2250	13	ADR28033 Human LIV
3	1359.5	85.0	2744	8	ACC50210 Breast ca
4	1359.5	85.0	2744	10	ADH28902 Human chr
5	1359.5	85.0	2744	12	ADN03826 Antipori

6	1359.5	85.0	2744	12	ADQ96115	Adq96115 T cell ac
7	1359.5	85.0	2744	13	ADR25539	Adr25539 Breast ca
8	1359.5	85.0	2744	13	ADR87269	Adr87269 CDNA enco
9	1359.5	85.0	2744	13	ADP56127	Adp56127 Human PRO
10	1359.5	85.0	2744	13	ADR66497	Adr66497 Human pro
11	1359.5	85.0	2744	13	ADR66155	Adr66155 Human pro
12	1359.5	85.0	3523	5	ABV25747	Abv25747 Human pro
13	1359.5	85.0	3537	11	ACN88857	Acn88857 Breast ca
14	1359.5	84.8	2268	8	ACC51044	Acc51044 Human Dia
15	1356.5	84.8	2268	12	ADL57160	Adl57160 Human NOV
16	1356.5	84.8	3461	6	ABK92205	Abk92205 Prostate
17	1356.5	84.8	3461	6	ABT07707	Abt07707 Breast ca
18	1356.5	84.8	3461	6	ABQ92299	Abq92299 Human bre
19	1356.5	84.8	3461	8	ABX76337	Abx76337 Lung can
20	1356.5	84.8	3461	11	ADN38737	Adn38737 Cancer/an
21	1356.5	84.8	3461	11	ADN39560	Adn39560 Cancer/an
22	1356.5	84.8	3461	11	ADN39520	Adn39520 Cancer/an
23	1355	84.7	1186	12	ADL57156	Adl57156 Human NOV
24	1352.5	84.6	2763	13	ACN40337	Acn40337 Tumour-as
25	1352.5	84.6	3445	6	ABS59534	Abs59534 Human CDN
26	1352.5	84.6	3445	12	ADH72211	Adh72211 Human gen
27	1352.5	84.6	3445	12	ADL57152	Adl57152 Human NOV
28	1351.5	84.5	2310	12	ADL57162	Adl57162 Human NOV
29	1349.5	84.4	2776	4	ADL13480	Adl13480 Human LIV
30	1347.5	84.3	3445	12	ADL57168	Adl57168 Human NOV
31	1345.5	84.1	3445	12	ADL57166	Adl57166 Human NOV
32	1345.5	84.1	3445	12	ADO39199	Ado39199 Human CDN
33	1340.5	83.8	3445	12	ADL56815	Adl56815 Human NOV
34	1324.5	82.8	2265	13	ADR87281	Adr87281 Liv-1 mut
35	1311.5	82.0	3461	4	ADL13479	Adl13479 Human LIV
36	1311.5	82.0	3586	10	ABT31919	Abt31919 Human bre
37	1265.5	79.1	2211	12	ADL57164	Adl57164 Human NOV
38	1189.5	74.4	3173	13	ACN43699	Acn43699 Human dia
39	1007	63.0	2236	12	ADN03825	Adn03825 Antipori
40	1007	63.0	2370	12	ADQ96253	Adq96253 T cell ac
41	959	60.0	1310	2	AAT33219	Aat33219 Oestrogen
42	959	60.0	1310	2	AAT99070	Aat99070 Partial s
43	959	60.0	2404	2	AAT33220	Aat33220 Oestrogen
44	959	60.0	2404	2	AAT99071	Aat99071 Partial s
45	875	54.7	1597	2	AAZ33622	Aaz33622 Human bre

ALIGNMENTS

RESULT 1	ADL57158	ADL57158 standard; DNA; 1101 BP.
XX	XX	XX
XX	ADL57158;	
XX	03-JUN-2004 (first entry)	
XX	Human NOV9c gene SEQ ID NO:103.	
XX	de; gene; human; antidiabetic; anorectic; cardiant; hypotensive;	
XX	antiartherosclerotic; anorectic; virucide; antibacterial; fungicide;	
XX	prozoacide; nootropic; neuroprotective; antiparkinsonian;	
XX	anticonvulsant; osteopathic; antiarthritic; antiinflammatory;	
XX	dermatological; antiaschemic; antipapemic; gene therapy;	
XX	fibroblast growth factor receptor 4; FGFR4;	
XX	complement factor I precursor; matrix metalloproteinase-15 precursor;	
XX	MDG3; T-lymphocyte surface antigen Ly-9 precursor;	
XX	fibroblast growth factor-21; FGF-21;	
XX	alpha-2 macroglobulin-like polypeptide variant;	
XX	antileukoprotease 1 precursor; LIV-1; nuclear hormone receptor NOR-1;	
XX	transmembrane protein-like; beta-neoendorphin-dynorphin precursor.	
OS	Homo sapiens.	
XX	XX	
XX	Key	Location/Qualifiers
XX	CDS	123..1031
XX	FT	/tag= a
XX	XX	

PN WO2004022723-A2.
 XX 18-MAR-2004.
 XX 09-SEP-2003; 2003WO-US028141.
 XX 09-SEP-2002; 2002US-0409145P.
 PR 10-SEP-2002; 2002US-0409544P.
 PR 12-SEP-2002; 2002US-0410320P.
 PR 16-SEP-2002; 2002US-0411060P.
 PR 23-SEP-2002; 2002US-0412766P.
 PR 24-SEP-2002; 2002US-0412825P.
 PR 25-SEP-2002; 2002US-0413342P.
 PR 30-SEP-2002; 2002US-0414832P.
 XX (CURA-) CURAGEN CORP.
 XX Zhong M, Guo X, Anderson DW, Ort T, Padigaru M, Rieger DK;
 PI WPI; 2004-315567/29.
 DR P-PSDB; ADL57159.
 XX New isolated NOVX polypeptides and polymucleotides, useful for
 PT preventing, diagnosing or treating NOVX-associated disorders, e.g.,
 PT osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
 PT asthma, or infections.
 XX Claim 17; SEQ ID NO 103; 214pp; English.
 PS The invention relates to a novel isolated polypeptide (NOVX) comprising a
 CC mature form of any of the 37 amino acid sequences fully defined in the
 CC specification. A polypeptide of the invention has antidiabetic,
 CC anorectic, antiobesity, hypotensive, antihypertensive, anorectic,
 CC virucidal, antibacterial, fungicidal, protozoacidal, nootropic,
 CC neuroprotective, antiparkinsonian, anticonvulsant, osteoprotective,
 CC antiarthritic, antiinflammatory, dermatological, antiseptic, and
 CC antileptemic activity. A polymucleotide of the invention may have a use
 CC in gene therapy. The polypeptides, nucleic acid molecules and antibodies
 CC are useful in the manufacture of a medicament for treating a syndrome
 CC associated with a human disease, preferably a NOVX-associated disorder.
 CC The nucleic acid molecules, polypeptides and antibodies are useful for
 CC treating, preventing or diagnosing diseases such as metabolic disorders,
 CC diabetes, obesity, infectious diseases (viral, bacterial, fungal,
 CC helminthic, and protozoal), anorexia, cancer, cardiovascular diseases
 CC (hypercholesterolemia, atherosclerosis), neurodegenerative disorders, Alzheimer's
 CC disease, Parkinson's disease, epilepsy, immune disorders
 CC (osteoarthritis), haematopoietic disorders, inflammatory skin disorders,
 CC asthma, and various dyslipidaemias. The nucleic acids and polypeptides
 CC may also be used as targets for the identification of small molecules
 CC that modulate or inhibit e.g. neurogenesis, cell differentiation, cell
 CC proliferation, haematopoiesis, wound healing and angiogenesis, in gene
 CC therapy. In generation of antibodies that bind immunospecifically to NOVX
 CC substances for use in therapeutic or diagnostic methods. The nucleic
 CC acids are further used as hybridisation probes, in chromosome mapping,
 CC tissue typing, preventive medicine, and pharmacogenomics. The NOVX
 CC polypeptides of the invention show homology to certain known human
 CC proteins: NOVXa-1e show homology to fibroblast growth factor receptor 4
 CC (FGFR4); NOVXa shows homology to complement factor I precursor; NOVXa
 CC shows homology to matrix metalloproteinase-15 precursor; NOVXa shows
 CC homology to MOC3; NOVXa-5c show homology to T-lymphocyte surface antigen
 CC LY-9 precursor; NOVXa-6m show homology to fibroblast growth factor-21
 CC (FGF-21); NOVXa-7c show homology to alpha-2 macroglobulin-like
 CC polypeptide variant; NOVXa-8g show homology to antileukoprotease 1
 CC precursor; NOVXa-9i show homology to LIV-1 protein; NOVXa shows homology
 CC to nuclear hormone receptor NOR-1; NOVXa-11j show homology to
 CC transmembrane protein-like; NOVXa-12c show homology to beta-neoendorphin
 CC -dynorphin precursor. The present sequence encodes a NOVX polypeptide of
 CC the invention.
 XX SQ Sequence 1101 BP, 283 A, 236 C, 262 G, 320 T, 0 U, 0 Other;

Alignment Scores:

Pred. No.:	2,31e-154	Length:	1101
Score:	1599.00	Matches:	302
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	12	Gaps:	0
US-10-659-004-104 (1-302) x ADL57158 (1-1101)			
QY	1	MetaIaArgLYLeuSerValIleLeuThrPheAlaLeuSerValThrAsnPro	20
DB	123	ATGGCGAGGAAGTATCTGTATCTTATCTTGAAGCTTGGCTCTGTGCACAAATCC	182
QY	21	LeuHISGLLeuLYsAlaAlaPheProGlnThrThrGluLYsLeuSerProAsnTrp	40
DB	183	CTTCATGAACCTMAAAGACGCTTCCCGACGACCACTGAGAAATTAGTCCGAAATTGG	242
QY	41	GLuSerGlyIleAsnValaAspLeuAlaIleSerThrArgGlnTrpHISLeuGlnLeu	60
DB	243	GAATCTGGCATTAATGTTGACTTGGCAATTTCCACGCGCAATTCATCTCAACAGCTT	302
QY	61	PheTYrArgTYrGlyGluAsnAsnSerLeuSerValGluGlyPheArgLYsLeuGln	80
DB	303	TTCTACCGCTATGAGAAATTAATTTCTTGTCACTTGAAGGCTTCAGAAATTAATCTTCA	362
QY	81	AsnIleGlyIleAspLYsIleLYsArgIleHISAlaHISAspHISAspHISAsp	100
DB	363	AATATAGCATAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA	422
QY	101	AspHISGLHISAspAspHISGLuArgHISAspHISGLuHISAspHISAspHIS	120
DB	423	GACACAGACATCACTGACACATGACGCTGACACATGACATGACATGACATGACAC	482
QY	121	HISProHISAspHISAspGlnArgTYrSerArgGluGluLeuLYsAspAlaGlyValAla	140
DB	483	CATCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCA	542
QY	141	ThrLeuAlaTrpMetValIleMetGlyAspGlyLeuHISAsnPheSerAspGlyLeuAla	160
DB	543	ACTTGGCTGAGATGGATATGATGGATGGATGGATGGATGGATGGATGGATGGATGG	602
QY	161	IleGlyAlaAlaPheThrGluGlyLeuSerSerGlyLeuSerThrSerValAlaValPhe	180
DB	603	ATTGCTGCTGCTTTTACTGAAGCTTATCAAGTGTATTAAGTATCTGTTGCTGCTG	662
QY	181	CysHISGLLeuProHISGLuLeuGlyAspPheAlaValLeuLeuLYsAlaGlyMetThr	200
DB	663	TGTCATGATGCTCTCATGAATTAAGTGAATTCCTGCTTCACTAAAGGCTGCGATGACC	722
QY	201	ValLYsGlnAlaValLeuLYsAsnAlaLeuSerAlaMetLeuAlaTYrLeuGlyMetAla	220
DB	723	GTTAAGCAGGCTGCTTATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA	782
QY	221	ThrGlyIlePheIleGlyHISTYrAlaGluAsnValSerMetTrpIlePheAlaLeuThr	240
DB	783	ACAGGAATTTTCACTTGGCATTAATGCTGAATAATTTCTTATGAGATTAATTTGCACTT	842
QY	241	AlaGlyLeuPheMetHISValAlaLeuValAspMetValProGluMetLeuHISAsnAsp	260
DB	843	GCTGCTTATTAACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	902
QY	261	AlaSerAspHISGLYsAspArgTYrGlyTYrPhePheLeuGlnAsnAlaGlyMetLeu	280
DB	903	GCTGATGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	962
QY	281	LeuGlyPheGlyIleMetLeuLeuIleSerIlePheGlnHISLYsIleValPheArgIle	300
DB	963	TTGGGTTTGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	1022
QY	301	AsnPhe 302	
DB	1023	AATTTTC 1028	

RESULT 2
ADR28033
ID ADR28033 standard; CDNA: 2250 BP.
XX
AC ADR28033;
XX
DT 04-NOV-2004 (first entry)
XX
DE Human LIV-1 protein encoding CDNA.
XX
KW Antibody; IGSF9; immunoglobulin superfamily member 9; LIV-1;
KW neoplastic disorder; vaccine; anti-idiotypic; cancer; antisense therapy;
KW cytoskeletal; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..2250
FT /tag= a
FT /product= "Human LIV-1 protein"
XX
XX MO200406933-A2.
XX
PD 12-AUG-2004.
XX
XX 27-JAN-2004; 2004WO-US002044.
XX
PR 27-JAN-2003; 2003US-0442535P.
XX
PA (MCLA/) MCLACHLAN K.
PA (GLAS/) GLASER S.
PA (PEAC/) PEACH R J.
PA (ROME/) ROME T.
XX
PI Mclachlan K, Glaser S, Peach RJ, Rowe T;
XX
XX WPI: 2004-580843/56.
XX P-PSDB: ADR28034.
XX
XX New antibody or its antigen binding fragment that associates with either
XX IGSF9 or LIV-1 polypeptide, and the polypeptides, useful in preparing a
XX vaccine composition for treating cancer.
XX
XX Disclosure; SEQ ID NO 28; 149pp; English.
XX
XX The invention relates to a novel isolated antibody or its antigen binding
XX fragment that associates with either IGSF9 (immunoglobulin superfamily
XX member 9) or LIV-1 polypeptide, and the fully defined polypeptides. The
XX polypeptide comprises a sequence selected from ADR28009, ADR28011,
XX ADR28027, ADR28028, ADR28030, ADR28031, and ADR28032, having
XX 966 or 979 amino acids. The invention further comprises a composition
XX comprising the antibody which associates with IGSF9 or LIV-1 for treating
XX a neoplastic disorder or the polypeptide comprising a sequence having 966
XX or 979 amino acids; a method of treating a mammal exhibiting a neoplastic
XX disorder; a vaccine for treating cancer comprising the IGSF9 or LIV-1
XX polypeptide or an anti-idiotypic antibody that immunologically mimics the
XX IGSF9 or LIV-1 antigen or its fragment and a carrier; a method of
XX inducing an immune response in a patient in need of treatment or
XX prevention of cancer; a method of diagnosing cancer by detecting
XX overexpression of IGSF9 or LIV-1 or its fragment; a method for
XX determining the prognosis of an individual receiving a cancer treatment;
XX a kit comprising the composition together with instructions for use to
XX treat or detect cancer; a method of treating a neoplastic disorder in a
XX mammal where the neoplastic cells express the IGSF9 or LIV-1 antigens; an
XX antisense nucleic acid up to 50 nucleotides in length comprising at least
XX an 8 nucleotide portion of IGSF9 or LIV-1 which inhibits the expression
XX of IGSF9 or LIV-1; a method of inhibiting the expression of IGSF9 or LIV-
XX 1 in cells or tissues; an isolated nucleic acid selected from the group
XX consisting of: ADR28008, ADR28010, ADR28017, ADR28018, ADR28019,
XX ADR28020, ADR28021, ADR28022, ADR28023, ADR28024, ADR28025, and ADR28026;
XX a vector comprising the nucleic acid; and a host cell comprising the
XX nucleic acid. The antibody has cytoskeletal activity. The antibody or its
XX antigen binding fragment that associates with either IGSF9 or LIV-1

CC polypeptide, and the IGSF9 or LIV-1 polypeptides, are useful in preparing
CC a composition for treating cancer. This polynucleotide sequence
CC represents the CDNA encoding a human LIV-1 protein of the invention.
XX
SQ Sequence 2250 BP; 657 A; 503 C; 474 G; 616 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 2,44e-129 Length: 2250
Score: 1359.50 Matches: 301
Percent Similarity: 40.32% Conservative: 1
Best Local Similarity: 40.19% Mismatches: 0
Query Match: 85.02% Indels: 447
DB: 13 Gaps: 1
US-10-659-004-104 (1-302) x ADR28033 (1-2250)
QY 1 MchlaarglylsleuservallleuulleuThrPhealaleuservallThrAsnPro 20
DB 1 ATGCGAGGAAAGTTATCTGTATCTTGCATCTGACCTTGGCCCTCTGTCAAAATCCC 60
QY 21 leuHiscgluleuylsalaalalaPheProGlnThrGluylsIleSerProAsnTrp 40
DB 61 CTTCACTGAAGCTAAAGACGCTGCTTCCCGACGACCTGAGAAATTAAGTCCGAATGG 120
QY 41 GluserglylleAsnValaAspleuAlaIleSerThrArgGlnTrhIleuGlnGlu 60
DB 121 GAATCTGGCAATTAATGTGACTTGGCAATTTCCACGCAATTCATCTCAACAGCTT 180
QY 61 PheTyArgTyrglygluAsnAsnSerleuservalluglyPheArglylsleuGln 80
DB 181 TTCTACCGCTATGAGAAATTAATCTTGTGCACTTGAAGGTTCAAGAAATTAATCTCA 240
QY 81 AsnIleglylleAspIlylleAspIlylleAspIlylleAspIlylleAspIlylle 100
DB 241 AATATAGGATAGATAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 300
QY 101 AspHisGluHisSerAspHisGluHisSerAspHisGluHisSerAspHisSer 119
DB 301 GACCAAGACATCACTGACGACGACGACGACGACGACGACGACGACGACGACGAC 360
QY 119 ----- 119
DB 361 GAGCATCACTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 420
QY 119 ----- 119
DB 421 CCAGACCATGATCAGATAGTTCAAGTAAAGATCTTAAGAACGAGGAGGAGAGACT 480
QY 119 ----- 119
DB 481 CACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 540
QY 119 ----- 119
DB 541 GTGACCTCACTGTGTACACACTGTCTGTGAAGAACTCACTTTTAGAGACAATAGAG 600
QY 119 ----- 119
DB 601 ACTCAAGACCTGGAAGAACTTCCCAAGATGTAAAGAGCTCACTCCACCAAGTGTG 660
QY 119 ----- 119
DB 661 ACATCAAGACCGGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGG 720
QY 119 ----- 119
DB 721 CCGGAAAGGCTTATGTATTCAGAAACCAATGAATTCCTCAGAGAGTGTTCAT 780
QY 119 ----- 119
DB 781 GCATCAAGCTATGACATCTCATGAGCATGGGATCCAGGTTCCGCTGAATGACAGAG 840
QY 119 ----- 119

Db	841	TTCAACTATCTCTGTCAGGCCATCATCAACCAATTGATGCTAGATCTTGTGATTCAT	900
Oy	119	-----	119
Db	901	ACAAAGTAAAAAGAGGTGAAATCCCTCCAAAGACCTATTTCATTAACAATAGCTGGTT	960
Oy	119	-----	119
Db	961	GGTGGTTTTATAGCCATTTCATTCATCAGTTTCTGTCTGTGCGGGGTATCTTAGTG	1020
Oy	119	-----	119
Db	1021	CCTCATGAATCGGGGTGTTTTTCAAAATTTCTCCGAGTTTCCTGTGTGGCACTGGCCGTT	1080
Oy	119	-----	119
Db	1081	GGGACTTTGAGTGTGATGCTTTTTTACACCTTCCTTCACATTCATGACAGTCAACAC	1140
Oy	119	-----	119
Db	1141	CATAGTCATAGCCATGAAGAACCAAGCAATGAAATGAAAAGGACCACTTTCAGTCAT	1200
Oy	119	-----	119
Db	1201	CTGTCTTCTCAAAACATAGAAAGAAAGTGCCATATTTGATTCAGGTGAAGGCTTAACA	1260
Oy	119	-----	119
Db	1261	GCTTAGAGAGCGCTGTATTTTCATGTTTTCTTTGGAACATGTCCTGACATTCATCAACAA	1320
Oy	119	-----	119
Db	1321	TTTAAAGATTAAGAAAGAAAAAGATCAGAAGAACTGAAATGATGATGATGGAGATT	1380
Oy	119	-----	119
Db	1381	AAGAAGCAATTGTCCAGATGATGATCTCACTTTCAACAAATAGAGAGAAATAGATACA	1440
Oy	119	-----	119
Db	1441	GATGATCGAATGAAAGGCTATTTCAGAGACATGACAAAGACCTCCCATTTGATTCCT	1500
Oy	119	-----	119
Db	1501	CAGCAGCCTGCAGTCTTGGAAAGAAAGAGTCATGATAGCTCATGCTCATCCACAGAA	1560
Oy	119	-----	119
Db	1561	GTCCTCAATGAATATGTACCCAGAGGTGCAGAAATTAATGCCATTTCACATTTCCAGAT	1620
Oy	119	-----	119
Db	1621	ACACTCGGCCAGTCAAGCATCTCATTCACCAATCATGACTACCATCATATTCTCCAT	1680
Oy	120	-----	133
Db	1681	CATCACCAACCAAAACCAACATCTCTCACTACACGCAACGCTACTCTGGAGAGAG	1740
Oy	134	LeuLYAAPAlagLYValAlaThrLeuAlaTrpMetValIImetGLYAAPGlyLeuHis	153
Db	1741	CTGAAAGATGCGCGCGTCGCCCACTTGGCCTGATGTGATAGGTGATGGCTGCACAC	1800
Oy	154	AanPheSerAspGlyLeuAlaIleGlyAlaAlaPheThrGluGlyLeuSerSerGlyLeu	173
Db	1801	AATTTCAAGAGAGGCTTAGCAAAATGGTGTGCTTTTACGAAAGGCTTATCAAGTGGTTTA	1860
Oy	174	SerThrSerValAlaValAPheCysHisGlyLeuProHisGlyLeuGlyAspPheAlaVal	193
Db	1861	AGTACTTCTGTTGCGTGTGTCTGTATGAGTTGGCTCAAGATTAAGTGACTTTGCTGT	1920
Oy	194	LeuLeuLYValagLYMetThrValIlyysGlnAlaValLeuTYrAsnAlaLeuSerAlaMet	213

Db	1921	CTACTAAAGCGTGGCAGTACCGCTTAAGCAGCGCTGTCCTTATATATGATTCGTACCGCATG	198
Qy	214	LeuAlaIATyLeuGIyMeIaIaIaThGIyLeIheIleGIyHiSTyAlaGIuAsnValSer	233
Db	1981	CTGGCGATATCTTGGAAATGGCAACAGGAATTTCTCATTTGGCTATTTATGGCAAAATGTTTCT	204
Qy	234	MeIcTrpIlePheAlaLeuThraIaGIyLeuPheMeThIsvAlaIaLeuValAspMeIVal	253
Db	2041	ATGGGATATATTTGCACTTACTGCTGCGCTTATTCATGATGTTGCTCTGGTGAATGGTA	2100
Qy	254	ProGIuMeIcLeuHiIaAsnAspAlaSerAspHiIaGIyCySerArgTrpGIyTyI-PhePhe	273
Db	2101	CCTGAATATGCTGCACAATGATGATGATGACCATGATGATGATGACCGCTGGGGGATTTCTTT	2160
Qy	274	LeuGIuIaIaIaGIyMeIcLeuLeuGIyPheGIyIleMeIcLeuLeuIleSerIlePheGIu	293
Db	2161	TTAAGAGATGCTGGGATGCTTTTGGGTTTGGAAATATGTTACTTATTTCCATATTGGAA	2220
Qy	294	HisIyIleValPheArgIleAsnPhe 302	
Db	2221	CATATAATCGTGTTTCGATATAATTTC 2247	
RESULT 3			
ACC50210			
ID	ACC50210	standard; cDNA; 2744 BP.	
AC	ACC50210;		
XX			
DT	12-JUN-2003	(first entry)	
DE	Breast cancer associated cDNA sequence SEQ ID NO:263.		
XX			
KW	Human; breast cancer; cytostatic; gene therapy; gene; ss.		
XX			
OS	Homo sapiens.		
XX			
FN	WO2003004989-A2.		
XX			
PD	16-JAN-2003.		
XX			
PF	21-JUN-2002; 2002WO-US019669.		
XX			
PR	21-JUN-2001; 2001US-0299887P.		
XX			
PR	27-JUN-2001; 2001US-0301572P.		
XX			
PR	18-JUL-2001; 2001US-0306501P.		
XX			
PR	25-SEP-2001; 2001US-0325002P.		
XX			
PR	05-MAR-2002; 2002US-0325285P.		
XX			
PR	14-MAY-2002; 2002US-0380391P.		
XX			
PA	(MILL-) MILLENIUM PHARM INC.		
XX			
PI	Lillie J, Gannavarapu M, Glatt K, Hoersh S, Kamatkar S;		
XX			
PI	Mertens M, Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE;		
XX			
PI	Baer RC, Hortobagyi GN, Puzstai L, Meric F, Sahin A, Mills GB;		
XX			
XX	WPI; 2003-210381/20.		
DR	P-F8DB; ABR47514.		
XX			
PT	Breast cancer diagnosis or treatment by comparing the level of expression		
XX			
PT	of a marker in a patient sample with that in the control non-breast		
XX			
PT	cancer sample.		
XX			
PS	Claim 1; SEQ ID NO 263; 128bp; English.		
XX			
XX			
CC	The present invention describes a method for assessing whether a patient		
XX			
CC	is afflicted with breast cancer. The method comprises comparing the level		
XX			
CC	of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and		
XX			
CC	ABR47386 to ABR47632) in a patient sample and the normal level of		
XX			
CC	expression of the marker in a control non-breast cancer sample, where a		
XX			
CC	significant increase in the level of expression of the marker in the		
XX			
CC	patient sample and the normal level is an indication that the patient is		
XX			
CC	afflicted with breast cancer. The breast cancer associated sequences from		
XX			
CC	the present invention have cytostatic activities and can be used in gene		

CC therapy. The method is useful for diagnosing and treating breast cancer.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WPO
 CC at ftp.wipo.int/pub/published_pat_sequences

XX Sequence 2744 BP; 786 A; 585 C; 601 G; 772 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3,19e-129	Length:	2744
Score:	1359.50	Matches:	301
Percent Similarity:	40.32%	Conservative:	1
Best Local Similarity:	40.19%	Mismatches:	0
Query Match:	85.02%	Indels:	447
DB:	8	Gaps:	1

US-10-659-004-104 (1-302) x ACC50210 (1-2744)

```

Qy 1 MetAlaArgLyLeuSerValIleuLeuThrPheAlaLeuSerValThrAspPro 20
Db 138 ATGGCGAGAAAGTTATCTGTATCTTGAATCCTGACCTTGGCCCTCTCTGTCACAAAATCC 197
Qy 21 LeuHiegluLeuValAlaAlaPheProGlnThrGluValIleSerProAsnTyr 40
Db 198 CTTGATGAATAAAGCAGCTGCTTCCCGACGACCACTGAGAAATTAATGTCGGAATTGG 257
Qy 41 GluSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTyrHisLeuGlnLeu 60
Db 258 GAATCTGGCATTTAATGTTGACTTGGCAATTTCCACGCGCAATTCATCTCAACAGCTT 317
Qy 61 PheTyrArgTyrGlyGluAsnAsnSerLeuSerValGluGlyPheArgLyLeuLeuGln 80
Db 318 TTCTACCCCTATGAGAAATTAATTTCTTGTCACTTGAAGGTTCAAGAAATTAATCTTCA 377
Qy 81 AsnIleGlyIleAspLyLeuValArgIleHisIleHisIleAspHisAspHisIleSer 100
Db 378 AATATAGGCATAGATTAAGATTAAGAAATCATATACCATGACGACGACCATCATCTCA 437
Qy 101 AspHisGluHisIleSerAspHisIleGluArgHisSerAspHisIleGluHisIleSer 119
Db 438 GACCACGAGCATCTCAGACCATGAGCGTCACTCAGACCATGAGCATCTCAGACCATC 497
Qy 119 ----- 119
Db 498 GAGCATCACTGACCATATCATATGCTGCTTCTGTGTAATAAAGCGAAAGCTTTTGC 557
Qy 119 ----- 119
Db 558 CCAAGCATGACTCAGATAGTTCAAGTAAAGATCTCTAGAAACAGCCAGGGGAAGAGCT 617
Qy 119 ----- 119
Db 618 CACCGACGAAACATGCGAGTGTAGAAAGATGTCAGAGCAGTGTAGTGTAA 677
Qy 119 ----- 119
Db 678 GTGACTCACTGTGTACACACTGTCTTGAAGAACTCACTTTAGAGACAATAGAG 737
Qy 119 ----- 119
Db 738 ACTCCAGAGACTGGAAAACTTCCCAAAAGATGTAAAGCAGCTCCACCCAGTGTG 797
Qy 119 ----- 119
Db 798 ACATCAAGAGCCGGGTGAGCCGGCTGGCTGTAGAAACAAATGAATCTGTGAGTAG 857
Qy 119 ----- 119
Db 858 CCCCCAAAAGGCTTTATGTATTTCCAGAAAACAATAAATCTCTCAGAGAGTGTTCAT 917
Qy 119 ----- 119
Db 918 GCATCAAGTACTGACATCTCATGGGATGGGATGCCAGTTCCGGTGAATGACAAAGAG 977

```

```

Qy 119 ----- 119
Db 978 TTCACTATCTGTGTCCAGCCATCATCAACCAATGTAGTGTGTTGATTCAT 1037
Qy 119 ----- 119
Db 1038 ACAAGTAAAGAGGCTGAATCCCTCCAAAGACTATTTCATTACAAATAGCTGGGTT 1097
Qy 119 ----- 119
Db 1098 GGTGTTTATAGCCATTTCATCATCATGATTTCTGTCTGTGCGGGTATCTTAGTG 1157
Qy 119 ----- 119
Db 1158 CTTCTCATGAATGGGGTGTTCATTAATTTCTCTGAGTTTCTTGGGACGTGGCGTT 1217
Qy 119 ----- 119
Db 1218 GGGACTTGAAGTGTGATGCTTTTACACCTTCTTCCACTTCATGCAAGTCACAC 1277
Qy 119 ----- 119
Db 1278 CATAGTCATAGCATGAAGAACAGCAATGGAATGAAGAGAGACACTTTTCAGTCAT 1337
Qy 119 ----- 119
Db 1338 CTGCTTCTCAAAACATAGAAAGAGGCTTATTTGATTCACAGTGAAGGTCATAA 1397
Qy 119 ----- 119
Db 1398 GCTTAGAGGCTGTATTTTCATGTTCTTGTGTAACATGCTCATTGATCAACAA 1457
Qy 119 ----- 119
Db 1458 TTTAAGATGAAGAAAAAATCAGAGAAACCTGAAATGATGATGTGAGATT 1517
Qy 119 ----- 119
Db 1518 AAGAGAGTTTCCAGTATGAATCTCACTTTCAACAAATGAGAGAAAGTAGATCA 1577
Qy 119 ----- 119
Db 1578 GATGATGAAGTGAAGCTATTTACAGACAGCTCAGAGGCTCCCACTTGAATCT 1637
Qy 119 ----- 119
Db 1638 CAGAGCTGCAAGCTTGAAAGAAAGAGGTATGATAGTCAATCCACAGAA 1697
Qy 119 ----- 119
Db 1698 GTCTACATGAATATGTACCCAGAGGTCGAGAAATGATCCATTCATTCACAGAT 1757
Qy 119 ----- 119
Db 1758 ACACTGGCCAGTCAGACATCTCATTCACACCATATGACTATCATATTTCTCAT 1817
Qy 120 ----- 119
Db 1818 CATCACCAACCAAAACCAATCTCCACAGTCAAGCCAGGCTCTCTCGGAGAGAG 1877
Qy 134 LeuLyAspAlaGlyValAlaThrLeuAlaTyrMetValIleMetGlyAspGlyLeuHis 1937
Db 1878 CTGAAGATCCCGGCTGCGACATTTGGCTGTGATGAATAGGGTGTATGGCTGCAC 1937
Qy 154 AspPheSerAspGlyLeuAlaIleGlyAlaAlaPheThrGluGlyLeuSerSerGlyLeu 173
Db 1938 AATTTACGCGATGGCTTACGCAATTTGGTGTCTTTTACCTAAGGCTTATCAAGGTTTA 1997
Qy 174 SerThrSerValAlaValPheCyHisIleGluLeuProHisIleGluGlyAspPheAlaVal 193
Db 1998 AGTACTTCTGTGTGTGTCTGTGATGAGTGGCTCATGAATTAAGTGAATTTGCTGTT 2057
Qy 194 LeuLeuLyAlaGlyMetThrValLyGlnAlaValLeuTyrAsnAlaLeuSerAlaMet 213

```

Db 2058 CTACTAAAGGCTGGCATATCCGTTNAGCAGGCTGTCCTTTATATGATTCAGCCATG 2117
 Qy 214 Leu1a1aTyrLeuGlyMetAlaThrGlyIlePheIleGlyHisTyrAlaGluAsnValSer 233
 Db 2118 CTGGCGATCTCTGGATATGGCAACAGGAAATTTTCATGTGTCATATGCTGGAATGTTCT 2177
 Qy 234 MetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisValAlaLeuValAspMetVal 253
 Db 2178 ATGTGATATATTGGCATCTTACCTGCTGGCTTATTCATGATATGTTGGCTGTGGATATGGTA 2237
 Qy 254 ProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSerArgTrpGlyTyrPhePhe 273
 Db 2238 CCTGAATATGCTGCACATGATGCTATGTAACCATGATGATGACCGCTGGGGGATTTCTTT 2299
 Qy 274 LeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeuLeuIleSerIlePheGlu 293
 Db 2298 TTACGAGATGCTGGATGCTCTTTTGGGTTTGGAAATATGTTACTATTTCCATATTGGA 2355
 Qy 294 HisIysIleValPheArgIleAsnDhe 302
 Db 2358 CATAAATCGTGTTCGTATTAATTTTC 2384
 RESULT 4
 ADH28902
 ID ADH28902 standard; DNA; 2744 BP.
 XX ADH28902;
 XX
 XX DT 11-MAR-2004 (first entry)
 XX
 XX DE Human chronic myelogenous leukaemia (CML) gene marker #170.
 XX
 XX KW ds; chronic phase chronic myelogenous leukaemia; CP-CML;
 XX KW blast crisis CML; BC-CML; human; chronic myelogenous leukaemia;
 XX KW gene marker.
 XX
 XX OS Homo sapiens.
 XX
 XX PN US2003104426-A1.
 XX
 XX PD 05-JUN-2003.
 XX
 XX PF 14-JUN-2002; 2002US-00171581.
 XX
 XX PR 18-JUN-2001; 2001US-0298914P.
 XX
 XX PA (LINSLEY LINSLEY P S.
 XX PA (MAOM/ MAO M.
 XX PA (DAIH/ DAI H.
 XX PA (HEYV/ HE Y.
 XX PA (RADICH/ RADICH J P.
 XX
 XX PI Linsley PS, Mao M, Dai H, He Y, Radich JP;
 XX
 XX DR WPI; 2003-787046/74.
 XX
 XX PT Classifying cell sample as chronic phase chronic myelogenous leukemia or
 XX PT blast crisis chronic myelogenous leukemia by detecting difference in
 XX PT expression of genes corresponding to the markers such as X15415, U89436.
 XX
 XX PS Disclosure; SEQ ID NO 170; 31pp; English.
 XX
 XX CC The invention relates to a method of classifying a cell sample as chronic
 XX CC phase chronic myelogenous leukemia (CP-CML) or blast crisis CML (BC-
 XX CC CML). The method is useful for classifying a sample as CP-CML or BC-CML.
 XX CC The present sequence represents a human chronic myelogenous leukemia
 XX CC (CML) gene marker used to distinguish blast crisis CML from chronic phase
 XX CC CML.
 XX
 XX SO Sequence 2744 BP; 786 A; 585 C; 601 G; 772 T; 0 U; 0 Other;
 Alignment Scores:

Pred. No.:	3,19e-129	Length:	2744
Score:	1359.50	Matches:	301
Percent Similarity:	40.32%	Conservative:	1
Best Local Similarity:	40.19%	Mismatches:	0
Query Match:	85.02%	Indels:	447
DB:	10	Gaps:	1
US-10-659-004-104 (1-302) x ADH28902 (1-2744)			
QY	1 MetalAaRgLySeuSeRvAlIleLeuIlleLeuThrPheAlaLeuSeRvAlThrAsnPro	20	
Db	138 ATGGCGAGGAAGTATCTGTAACTTGAATCTTGAACCTTGGCCCTCTGTGCACAAATCC	197	
QY	21 LeuHlsgLueuLySaIaAlaPheProGlnThrThrguLySileSerProAsnTrp	40	
Db	198 CTTATGAACATAAAGACGCTGCTTCCCGACGCCACTGAGAAATTAGTCCGAATTGG	257	
QY	41 GluSeRgLyIleAsnValAspLeuAlaIleSerThraRgInTrhIleuGlnLeu	60	
Db	258 GAATCTGGCATTAATGTGACTTGGCAATTTCCACAGGCAATATCATCTCAACAGCTT	317	
QY	61 PheTyraRgTyRgLyGluAsnAsnSerLeuSeRvAlGluGlyPheAlaRgLySeuLeuGln	80	
Db	318 TTCTACCGCTATGAGAAATAATCTTGTTCACAGTTGAAGGGTTCAGAAATTAATTCTTCA	377	
QY	81 AsnIleGlyIleAspLySileLyAsRgIlleHisIleHisAspHisAspHisSer	100	
Db	378 AATATAGCATAGATTAAGATTAAAGATTCATATACCATGACACGACCATCATCTCA	437	
QY	101 AspHisGluHisIleSerAspHisGluArgHisSerAspHisGluHisIleSerAsp	119	
Db	438 GACCAACGAGCTACTCAGACACAGAGGTCACTCAGACCATGACATCACTCAGACAC	497	
QY	119 -----	119	
Db	498 GAGCATCACTGACCAATATCATGCTGCTTGTGTAATAAGCGAAAGCTCTTTC	557	
QY	119 -----	119	
Db	558 CCAAGCATGACTAGATAGTTCAAGTAAAGATCTTAGAAACAGCCAGGAGAAAGAGCT	617	
QY	119 -----	119	
Db	618 CACCGACCAAACTGCCAGGTGAGAAAGAAATGCAAGACAGTGTAGTGTGA	677	
QY	119 -----	119	
Db	678 GTGACCTCAACTGTGTACAACTGTCTCTGAGAAACTCACTTTCTAGACAATAGAG	737	
QY	119 -----	119	
Db	738 ACTCCAAAGCTGGAAAATCTTCTCCCAAGATGTAGACAGCTCCATCCACCCAGTGT	797	
QY	119 -----	119	
Db	798 ACATCAAAAGACCGGGGTAGCCGGCTGGCTGGTGTAGAAAACAATGATCTGTGATGAG	857	
QY	119 -----	119	
Db	858 CCCCAGAAAGCTTATGTATTCAGAAACAATAAGAAAATCCTCAGAGAGTTTCAAT	917	
QY	119 -----	119	
Db	918 GCATCAAGCTATGACATCTCATGGCATGGGCATCCAGTTCCGCTGAATGCAACAGAG	977	
QY	119 -----	119	
Db	978 TTCAACTATCTGTGTGACCAATCATCAACAATTTGATGCTAGATCTTGTGTGATTCAT	103	
QY	119 -----	119	
Db	1038 ACAATGAAAAGAGGTGAATCCCTCCAAAGACTTATTCATTAACAATAGCTGGTT	1097	

```

QY 119 ----- 119
DB 1098 GGTGTTTATAGCATTTCCATCATCAGTTTCCTGCTCTGCGGGGTTATTTAGTG 1157
QY 119 ----- 119
DB 1158 CCTCATGAAATCGGGGTGTTTTCAAAATTTCTCTGAGTTTCCTTGCGCACTGCGGTT 1217
QY 119 ----- 119
DB 1218 GGGACTTTGAGTGATGCTTTTTCACCTTTCCATTCATGCAAGTCACGAC 1277
QY 119 ----- 119
DB 1278 CATAGTCATAGCATGAAGAACCCAGCAATGGAATGAAAGAGACCACTTTTCAGTCAT 1337
QY 119 ----- 119
DB 1338 CTGCTTCTCAAAACATAGAAAGAGTGCTATTTTGAATTCACAGTGGAAGGCTTAACA 1397
QY 119 ----- 119
DB 1398 GCTCTAGAGGCGCTGATTTTCATGTTCTGTTGAACATGTCCTCATGATGATCAACAA 1457
QY 119 ----- 119
DB 1458 TTTAAAGATGAAGAAAAGAAATCAGAAAGAACTGAAAATGATGATGATGAGATT 1517
QY 119 ----- 119
DB 1518 AAGAAGCAGTTGTCGAATGATGATCTCACTTTCAACAAATGAGAGAAAGTAGATACA 1577
QY 119 ----- 119
DB 1578 GATGATCGAAGTGAAGGCTATTTACGAGACAGTCAAGAGCCCTCCCACTTGATTTCT 1637
QY 119 ----- 119
DB 1638 CAGCAGCTGCAAGTTTGGAAGAGAGAGAGTATGATGATGATGATGATGATGATGAT 1697
QY 119 ----- 119
DB 1698 GTTACATGAATATGATACCCAGAGGTCGAAGAAATGATGATGATGATGATGATGATGAT 1757
QY 119 ----- 119
DB 1758 ACACCTGCGCAGTACAGCATCTCATCCACCATGATGATGATGATGATGATGATGATGAT 1817
QY 120 ----- 119
DB 1818 CATCACCAACCAAAACCAACCATCTCATGATGATGATGATGATGATGATGATGATGATGAT 1877
QY 134 LeuLYSAPALAGLYValAlaThrLeuAlaTrpMetValIleMetGlyAspGlyLeuHis 153
DB 1878 CTGAAAGATGCGGGGTCGCCACTTTGGCTGCGATGATGATGATGATGATGATGATGATGAT 1937
QY 154 AenPheSerArgGlyLeuAlaIleGlyValAlaPheTrpGlyLeuSerSerGlyLeu 173
DB 1938 AATTTACAGCATGAGCTAGCAATGCTGCTCTTATCTGAAGGCTTATCAAGTGGTTA 1997
QY 174 SerThrSerValAlaValAlaPheCysHisIleLeuProHisIleGlyLeuAspPheAlaVal 193
DB 1998 AGTACTTCTGTTGCTGTTCTGTCATAGATGATGATGATGATGATGATGATGATGATGATGAT 2057
QY 194 LeuLeuLYSALAGLYMetThrValIleGlyValAlaValIleuTYRAsnAlaLeuSerAlaMet 213
DB 2058 CTACTAAAGGCTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2117
QY 214 LeuAlaTYRLeuGlyMetAlaThrGlyIlePheIleGlyHisIleTYRAlaGlyAsnValSer 233
DB 2118 CTGGCGATCTTGTGAATGCAACAGAAATTTTCATGATGATGATGATGATGATGATGATGAT 2177
QY 234 MetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisValAlaLeuValAspMetVal 253

```

```

DB 2178 ATGTGAAATTTGCACTTACGCTGCGCTTATTCATGATGATGATGATGATGATGATGAT 2237
QY 254 ProGluMetLeuHisAenAspAlaSerAspHisGlyCysSerArgTrpGlyTYRPhePhe 273
DB 2238 CCTGAAATGCTGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2297
QY 274 LeuGlnAenAlaGlyMetLeuGlyPheGlyIleMetLeuAlaIleSerIlePheGly 293
DB 2298 TTACAGAAATGCTGCGATGCTTTGGGTTTGGAATGATGATGATGATGATGATGATGATGAT 2357
QY 294 HisIleValAlaPheArgIleAsnPhe 302
DB 2358 CATAAATGCTGTTTCGATTAATTTTC 2384

RESULT 5
ADN03826
ID ADN03826 standard; cDNA; 2744 BP.
XX
AC ADN03826;
XX
DT 01-JUL-2004 (first entry)
XX
DE Antipsoriatic cDNA sequence #112.
XX
KM ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS Homo sapiens.
XX
PN MO2004028479-A2.
XX
PD 08-APR-2004.
XX
PF 25-SEP-2003; 2003WO-US030907.
XX
PR 25-SEP-2002; 2002US-0414006P.
XX
PA (GETH ) GENENTECH INC.
XX
PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
PI Wu TD;
XX
DR WPI: 2004-305105/28.
DR P-PSDB; ADN03827.
XX
PT New PRO nucleic acid or polypeptide, useful for preparing a
PT pharmaceutical composition for diagnosing or treating psoriasis in a
PT mammal.
XX
PS Claim 1; SEQ ID NO 220; 3069bp; English.
XX
CC The invention relates to novel polynucleotide and polypeptides for
CC treating psoriasis or a sequence having at least 80% identity to the
CC above sequences. The nucleic acid is useful for preparing a composition
CC for diagnosing or treating psoriasis in a mammal. This sequence
CC corresponds to one of the polynucleotides of the invention.
XX
SQ Sequence 2744 BP; 786 A; 585 C; 601 G; 772 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3,18e-129 Length: 2744
Score: 1359.50 Matches: 301
Percent Similarity: 40.32% Conservative: 1
Best Local Similarity: 40.19% Mismatches: 0
Query Match: 85.02% Indels: 447
DB: 12 Gaps: 1

US-10-659-004-104 (1-302) x ADN03826 (1-2744)
QY 1 MetAlaArgLYSLeuSerValIleLeuIleuThrPheAlaLeuSerValThrAsnPro 20
DB 138 ATGGGAGGAAGATTATCTGATATCTTGAATCTTGCCTTGGCCCTCTGTCACAAATGCC 197

```

QY	21	LeuHi8GluLeuLyAlaAlaAlaPheProGlnThrTrgGluLySIIeserProAsnTrp	40
Db	198	CTTCATGAACTAAAGACAGCTGCTTCCCCGAGCACCTGAGAAATTAAGTCCAAATGG	257
QY	41	GluSerGlyIleAsnValAspLeuAlaIleSerThrAspGlnTrpGHIslEugInGlnLeu	60
Db	258	GAATCTGGCATTAATGTTGACTTGGCAATTTCCACGGCAATTCATCTACACAGCTT	317
QY	61	PheTyArgTyrgIyGluAsnAsnSerLeuSerValGluGlyPheArgLySLeuLeuGln	80
Db	318	TTTCACCGCATATGAGAAATAATATCTTTGTCAGTTGAAGGGTTCAAGAAATTACTTCAA	377
QY	81	AsnIIegIyIleAspLySIIelybArgIIehiSIIehiShIAspHIAspHIHiser	100
Db	378	AAATATAGGCATATGATTAAGATTAAAGAAATCCATATACACATGACCAAGCATCACTCA	437
QY	101	AspHI8GluHI8HiserAspHI8GluAspHI8GluAspHI8GluHI8HiserAsp---	119
Db	438	GACACGAGCATCACTCAAGCACATGAGCTCACTCAACATGAGCATCACTCAAGACCAC	497
QY	119	-----	119
Db	498	GAGATCACTCTGACCATATATCATGCTGCTTGTGGTAAATAATAGCGAAAGCTCTTGC	557
QY	119	-----	119
Db	558	CCAGACCATGATCAATAGATTTCAGGTAAGATCCTAGAAACAGCAGGGAAAGAGCT	617
QY	119	-----	119
Db	618	CACCGACAGACATGCCAGTGGTGAGAGAAATGTCAGACAGTGTTAAGTCTAGTAA	677
QY	119	-----	119
Db	678	GTGACTCACTGTGTATCAACAACCTGTCTGTGAAGAACTCACTTCTTAGAGACAATAGAG	737
QY	119	-----	119
Db	738	ACTCCAAAGACTGGAAACTCTTCCCAAGATGTAAAGACTCCACTCCACCAAGTTC	797
QY	119	-----	119
Db	798	ACATCAAAAGCCGGGTGAGCCGGCTGGCTGTAGAAAACAAATGAATCTGTGATGAG	857
QY	119	-----	119
Db	858	CCCCGAAAGGCTTTATGTATTCAGAAACAAATGAAGAAATCCTCAGAGTGTTCAT	917
QY	119	-----	119
Db	918	GCATCAAGACTAGCATATCTCATGGCATGGGCATCCAGTTCGCTGAATGCAACAGAG	977
QY	119	-----	119
Db	978	TTCAACTATCTGTGCCAGCATATCAACCAATGTATGCTAGATCTTGTGTATTCAT	1037
QY	119	-----	119
Db	1038	ACAAGTAAAGAAAGGCTGAATCCTCCAAAGACTTATTCATTACAAATAGCTGGGTT	1097
QY	119	-----	119
Db	1098	GGTGTATTATAGCCATTCATCATCAGTTTCTGTCTGTGCGGGGTTATCTTAGTG	1157
QY	119	-----	119
Db	1158	CTCTCATGAATCGGGTGTTCCTCAATTCCTCTGAGTTTCCTGTGGCACTGGCCGT	1217
QY	119	-----	119
Db	1218	GGGACTTTAGTGGTATGCTTTTTCACCTTTCCACATTCATGACAGTCAACAC	1277
QY	119	-----	119

Db	1278	CATAGTCATAGCCATGAAAGAACGACGATGAGAAATGAAAAGAGACCATTTTCAGTCAT	133
QY	119	-----	119
Db	1338	CTGTCTTCTAAAACATAGAGAAAGTCCATTTTGTATTCACGTGAGAGGTCTAAC	1397
QY	119	-----	119
Db	1398	GCTCTAGAGAGCCTGTATTTCATGTTTCTTGTTGAACATGTCCTCACATTGATCAACAA	1457
QY	119	-----	119
Db	1458	TTTAAAGATAGAAAGAAAAGATCAGAGAAACCTGAAAATGATGATGTGAGATT	1517
QY	119	-----	119
Db	1518	AAGAAGCAGTTGTCCAGTAGTATGATCTCAACTTTCACAAATGAGAGAAAGTAGATACA	1577
QY	119	-----	119
Db	1578	GATGATCGACTGAGGCTATTTCAGAGACACTCACAGAGCCCTCCACTTGTATCT	1637
QY	119	-----	119
Db	1638	CAGCAGCCTGCAGTCTTGAGAAAGAAAGAGTCAATGATGCTCATGCTCATCCACAGAA	1697
QY	119	-----	119
Db	1698	GTTCAATGATATGTATGTACCCAGAGGGTGCAAGATAAATGCCATTCACATTTCCACGAT	1757
QY	119	-----	119
Db	1758	ACACTGGCCAGTCAGACGATCTCATTCACACATCATGATCTTCATCATATTCTCCAT	1817
QY	120	-----HisHisProHisSerHisSerGlnArgTyrSerArgGlnGlu	133
Db	1818	CATCACACACACAAAACCAACCATCTCTCACAGTCACAGCCAGCGCTACTCTCGGAGAG	1877
QY	134	LeuLysAspAlaGlyValAlaThrLeuAlaTrpMetValIleMetGlyAspGlyLeuHis	153
Db	1878	CTGAAAGATGCCGGCGTGCACCATTTGGCCCTGATGGATTAATGGGATGAGCGCTGCAC	1937
QY	154	AsnPheSerAspGlyLeuAlaIleGlyValAlaIlePheThrArgGlnGlyLeuSerSerGlyLeu	173
Db	1938	AATTTCAACGATGGCCCTACGCAATTGGTCTCTTTACTGAAGCTTATCAAGTGGTTA	1997
QY	174	SerThrSerValAlaValPheCysHisGlyLeuProHisGlyLeuGlyAspPheAlaVal	193
Db	1998	AGTACTTCTGTGCTGTGTCTGTGCATGAGTGTGCTCATGAATTAGGTGACTTGTCTGTT	2057
QY	194	LeuLeuLysValAlaGlyMetThrValLysGlnAlaValLeuTyrAsnAlaLeuSerAlaMet	213
Db	2058	CTACTAAAGGCGTGCATACGACGTTTAAGAGGCTGTCTTTAAATGCAATTGTCCAGCCATG	2117
QY	214	LeuAlaTyrLeuGlyMetAlaThrGlyIlePheIleGlyHisTyrAlaGlnAsnValSer	233
Db	2118	CTGGCGTATCTTGGAATGGCAACAGAAATTTTCATTGGTCAATTATGCTGAATACTTTCT	2177
QY	234	MetTyrIlePheAlaLeuThrAlaGlyLeuPheMetHisValAlaLeuValAspMetVal	253
Db	2178	ATGTGATATATTGCACTTACTGTGCGCTTATTTCATGTATGTGTGCTCTGTTGATATGTGTA	2233
QY	254	ProGlnMetLeuHisAsnAspAlaSerAspHisGlyCysSerArgTrpGlyTyrPhePhe	273
Db	2238	CCTGAAAGCGTGCACAAATGATGCTAGTACCATGATGATGTACCCGCTGGGGGATTTCTTT	2297
QY	274	LeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeuLeuIleSerIlePheGlu	293
Db	2298	TTACGAAATGCTGGGATGCTTTTGGGTTTGGATTATGTAATTACTTATTTCATATTGAA	2357
QY	294	HisLysIleValPheArgIleAsnPhe	302


```

QY 119 ----- 119
Db 1218 GGGACTTTGAGTGGATGCTTTTACACCTTCTTCCATCTTCAGCAAGTCAACAC 1277
QY 119 ----- 119
Db 1278 CATAGTCATAGCCATGAGAACACAGCAATGAAATGAAAAAGAGACCACTTTTCACTCAT 1337
QY 119 ----- 119
Db 1338 CTGTCTTCTCAAAACATGAGAGAAAGTGCCTATTTTGAATTCACGTCGAGAGGCTTAACA 1397
QY 119 ----- 119
Db 1398 GCTTAGAGAGCCCTGATTTTCATGTTTCTTGTGAACATGCTCCTCATGATGATCAACAA 1457
QY 119 ----- 119
Db 1458 TTTAAAGATAGAGAGAAAAAGATCAGAGAAACCTGAAAATGATGATGATGAGATTT 1517
QY 119 ----- 119
Db 1518 AAGAAAGAGTGTCCAGATGATGATCTCACTTCAACAAATGAGAGAAAGTAGATACA 1577
QY 119 ----- 119
Db 1578 GATGATCGAAGTGAAGGCTATTTCAGAGACACTCAGAGAGCCCTCCCATTTGATTTCT 1637
QY 119 ----- 119
Db 1638 CAGAGCCTGACAGTCTTGAGAGAGAGAGATCATGATAGCTCATGCTCATCCAGAGAA 1697
QY 119 ----- 119
Db 1698 GTCTCAATGAATATGTATCCAGAGGTCGAGAAATTAATGCCATTCACTTTCCAGCAT 1757
QY 119 ----- 119
Db 1758 ACACCTGGCCATCAGACGATCTCATTCACACCATCATGATACCATCATATTTCCAT 1817
QY 120 ----- 119
Db 1818 CATCAACCAACCAAAACCAACCATCTCCTCAGCTCAGCAGCCGCTACTCTGGAGAGAG 1877
QY 134 LeuLYsAspAlaGlyValAlaThrLeuAlaTrpMetValIleMetGlyAspGlyLeuHis 153
Db 1878 CTGAAGAAGATGCGCGGCTGCGCCTTGGCTGGATGGATATATGAGTGGCTGAC 1937
QY 154 AsnProSerAspGlyLeuAlaIleGlyAlaAlaPheThrGluGlyLeuSerSerGlyLeu 173
Db 1938 AATTTCAAGGATGGCTAGCATGTGTGCTTTTACTGAAAGGCTTATCAAGTGGTTA 1997
QY 174 SerThrSerValAlaValAlaPheCysHisGlyLeuProHisGlyLeuGlyAspPheAlaVal 193
Db 1998 AGTACTTCTGTGTGCTGTGTTCTGTATGATGATGCTCTCATGATTAAGGACCTTGTCTGT 2057
QY 194 LeuLeuLYsAlaGlyMetThrValLYsGlnAlaValLeuLYsAspAlaLeuSerAlaMet 213
Db 2058 CTACTAAAGAGCTGGCATGACCCCTTAAGCAGGCTGCTCTTATTAAGCATTTGACACCCATG 2117
QY 214 LeuAlaLYsLeuGlyMetAlaThrGlyIlePheIleGlyHisLYsLYsAlaGlyAsnValSer 233
Db 2118 CTGGCGGATCTTGGATGGCAACAGGAATTTTCATTGGTCATTATGCGTGAATGTTTTCT 2177
QY 234 MetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisValAlaLeuValAspMetVal 253
Db 2178 ATGTGGAATATTGCACTTACTCTGCTTATTCATGATGATGCTCTGCTGATATGTA 2237
QY 254 ProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSerSerArgTrpGlyTyrPhePhe 273
Db 2238 CCTGAATGCTGCACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2297
QY 274 LeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeuLeuIleSerIlePheGlu 293

```

```

Db 2238 TTAAGAAATGCTGGAGATGCTTTTGGTTTGGATTTATGTTACTTATTTCCATATTTGAA 2357
QY 294 HisLYsIleValIlePheArgIleAsnPro 302
Db 2358 CATAAATCGTGTTCGTATTAATTTTC 2384

RESULT 7
ID ADR25539 standard; DNA; 2744 BP.
AC ADR25539;
DT 21-OCT-2004 (first entry)
DE Breast cancer prognosis marker #1400.
ds; breast cancer; prognosis; gene expression; diagnosis.
OS Homo sapiens.
PN W02004065545-A2.
PD 05-AUG-2004.
PE 15-JAN-2004; 2004WO-US001100.
PR 15-JAN-2003; 2003US-00342887.
PA (ROSE-) ROSETTA INPHARMATICS LLC.
PA (NECA-) NETHERLANDS CANCER INST.
PI Van't Veer LJ, He Y;
WPI; 2004-593473/57.

Classifying a breast cancer patient according to prognosis comprises
determining the similarity between the level of expression of each of
five genes in a cell sample taken from patient, to control levels.

Disclousure; SEQ ID NO 1400; 226bp; English.

The invention relates to a method of classifying a breast cancer patient
according to prognosis by determining the similarity between the level of
expression of each of five genes for which markers are listed in the
CC specification, in a cell sample taken from the breast cancer patient, to
control levels of expression for each respective five genes to obtain a
CC patient similarity value. The methods are useful for classifying a breast
CC cancer patient according to prognosis. Kits and computer program products
CC are useful for data analysis using the diagnostic, prognostic and
CC statistical methods of the invention. This sequence corresponds to a
CC marker used in the method of the invention.

XX
SQ Sequence 2744 BP; 786 A; 585 C; 601 G; 772 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3 139-129 Length: 2744
Score: 1359.50 Matches: 301
Percent Similarity: 40.32% Conservative: 1
Best Local Similarity: 40.19% Mismatches: 0
Query Match: 85.02% Indels: 447
DB: 13 Gaps: 1

US-10-659-004-104 (1-302) X ADR25539 (1-2744)
QY 1 MetAlaArgLYsLeuSerValIleLeuIleLeuThrPheAlaLeuSerValThrAsnPro 20
Db 138 ATGGCGAGGAAGTATATCTGATCTGATCTGACCTTTGCCCTCTGTCACAAATCCC 197
QY 21 LeuHisGluLeuLYsAlaAlaAlaPheProGlnThrThrGluLYsIleSerProAsnTrp 40
Db 198 CTTCATGAACTTAAGAGAGCTGCTTTCCCCAGACCACTGAGAAATTAATGATCCGAATTTGG 257

```

Qy 41 GluSerGylIleAsnValAspLeuValIleSerThrArgIleTyrHisLeuGlnLeu 60
Db 258 GAATGGCATTTAATGTTGACTTGACATTTCCACACGCGCATATCATCTCAACAGCTT 317
Qy 61 PheTyrArgIleGluAsnAsnSerLeuSerValGluGlyPheArgLeuLeuGln 80
Db 318 TTCTACCGCTATGGAGAAATATATCTTGTGACGTGAAGGTTCAAGAAATTACTTCAA 377
Qy 81 AsnIleGlyIleAspLeuValIleYerArgIleHisGlyHisIleAspHisIleSer 100
Db 378 AATATAGCATATGATATGATTAAGAAATCCATATACACATGACACGACCATCATCTCA 437
Qy 101 AspHisGluHisIleSerAspHisGluArgHisIleSerAspHisIleSerAsp 119
Db 438 GACCAAGCATCTCAAGACCATGAGCGTCACTCAGACATGACATCATCTCAGACAC 497
Qy 119 ----- 119
Db 498 GAGATCACTCGACCATATCATGCTGCTTCTGTATAAAATPAGCGAAAGCTTTTGC 557
Qy 119 ----- 119
Db 558 CCAAGCATGATGATAGTATGATTAAGATCCTAGAAACAGCGGAGGAAAGAGCT 617
Qy 119 ----- 119
Db 618 CACCGACCAAGAACATGCCAGTGTAGAAAGAAATGTCAGACAGTGTATGCTAGTGA 677
Qy 119 ----- 119
Db 678 GTGACCTCACTGTGTACACAGCTGTCTGTAGAGAACTCACTTTCTAGAGACAATAGAG 737
Qy 119 ----- 119
Db 738 ACTCGAAGCCTGGAAAACTCTCCCAAGATGTAGACAGCTCCACCCAGTGTCTC 797
Qy 119 ----- 119
Db 798 ACATCAAAAGACCGGGGTAGCCGGCTGGCTGTAGGAAAACAAAATGAAATCTGTAGTGA 857
Qy 119 ----- 119
Db 858 CCGCGAAAAAGGCTTATGTATTCAGAAAAACAATAATGAATCTCAGAGATGTTTCAT 917
Qy 119 ----- 119
Db 918 GCATCAAGCTACTGACATCTCATGGCATGGGATCCAGTTCGCTGATGATCAACAGAG 977
Qy 119 ----- 119
Db 978 TTCAACTATCTCTGTCCAGCATCATCAACCAAAATGTAGTGAATCTTGTGTGATTCAT 1037
Qy 119 ----- 119
Db 1038 ACAAGTGAAGAAGGCTGAATCCCTCCAAAGCTTATTCATTAACAATAGCCTGGCTT 1097
Qy 119 ----- 119
Db 1098 GGTGTTTTATAGCATTTTCATCATCAGTTTCCGTCTGTCTGTGGGGTATCTTAGTG 1157
Qy 119 ----- 119
Db 1158 CCTCTCATGAATCGGGTGTCTTTCAAATTTCTCTGAGTTTCTTGTGGCACTGGCGCTT 1217
Qy 119 ----- 119
Db 1218 GGGACTTGTAGTGTATCTTTTACACCTTCTTCACATTCATGCAAGTCAACAC 1277
Qy 119 ----- 119
Db 1278 CATAGTCATAGCATGAGAAACAGCAATGGAATGAAGAGACCACTTTTCAGTTCAT 1337
Qy 119 ----- 119

Db 1338 CTGTCTTCTCAAAACATAGAAAGAAAGTCCATTTTGTGATTCAGAGTGAAGGCTTAACA 1397
Qy 119 ----- 119
Db 1398 GCTCTAGAGAGCCTGTATTTTCATGTTTCTTGTGAACATGCTCATGATGATCAACAA 1457
Qy 119 ----- 119
Db 1458 TTTAAGATAGAAAGAAAGAAATCAGAGAAACCTGAAATGATGATGTGAGATT 1517
Qy 119 ----- 119
Db 1518 AAGAGACGTTGCTCAAGTATGAATCTCACTTTCAACAAATGAGAGAAAGATGATACA 1577
Qy 119 ----- 119
Db 1578 GATGATGAACTGAAGGCTATTTAGAGACAGACTCACAGAGCCCTCCACTTTGATTCT 1637
Qy 119 ----- 119
Db 1638 CAGAGCCTGCACTTTGAGAGAAAGAGTCAATGATAGCTCATGCTCATCCACAGAA 1697
Qy 119 ----- 119
Db 1698 GTCTAATGAATATGTACCAGAGGTGCAAGATTAATGCCATTCAATTTCCAGAT 1757
Qy 119 ----- 119
Db 1758 ACATCGGCGCAGTACAGATCTCATTCACCAACCATCATGACTCATATTTCTCCAT 1817
Qy 120 ----- 1817
Db 1818 CATGACCAACCAAAACCAACATCTCAGAGTACAGCCGCGCTACTCTCGGAGAGAG 1877
Qy 134 LeuValAspAlaGlyValAlaThrLeuAlaTyrMetValIleMetGlyAspGlyLeuHis 1937
Db 1878 CTGAAGATGCCGCGTCCGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 1937
Qy 154 AsnPheSerAspGlyLeuAlaIleGlyAlaAlaPheThrArgIleLeuSerSerGlyLeu 173
Db 1938 AATTTCAAGCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1997
Qy 174 SerThrSerValAlaValPheCysHisGlyLeuProHisGlyLeuGlyAspPheAlaVal 193
Db 1998 AGTACTTCTGTGCTGTGCTGTGCTGATGATGATGATGATGATGATGATGATGATGAT 2057
Qy 194 LeuLeuValAlaGlyMetThrValLeuGlyAlaValLeuTyrAsnAlaLeuSerAlaMet 213
Db 2058 CTACTAAGGCTGCGCATGACCGTTAGCAGAGCTGTCTTATATATGATGATGATGATGAT 2117
Qy 214 LeuAlaTyrLeuGlyMetAlaThrGlyIlePheIleGlyHisTyrAlaGluAsnValSer 233
Db 2118 CTGGCGTATCTTGGAAATGGCAACAGAAATTTTCAATTTGATGATGATGATGATGATGAT 2177
Qy 234 MetTyrIlePheAlaLeuThrAlaGlyLeuPheMetHisValAlaLeuValAspMetVal 253
Db 2178 ATGTGATATTTGACATTAAGCTGCTGTATTCATGATGATGATGATGATGATGATGAT 2237
Qy 254 ProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSerArgTyrPhePhe 273
Db 2238 CTTGAATGCTGCAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2297
Qy 274 LeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeuLeuIleSerIlePheGly 293
Db 2298 TTACAGAAATGCTGGATGCTTTTGGTGTGAAATATGTTACTTATTTTCATATTTGAA 2357
Qy 294 HisIleValIlePheArgIleAsnPhe 302
Db 2358 CATAAATCGTGTTCATATAAATTC 2384

RESULT 8
ADR87269

[illegible]

Dd	2238	CCTGAAAGTCGCACCAATGAATGCTAAGTACCATGGATTACCGGTGGAGTAATTTCTT	2297
Oy	274	LeuglnaenAlaGLymetleuleuGlYphegilylmetleuleuileSerilePheglu	293
Dd	2236	TTCAGAAATCCTGGAGATCCITTTGGGTTTTTGGAATTACTTACTTATTTCCATAATTGAA	235
Oy	294	HlslyslIvalIpheargIlasaphe	302
Dd	2358	CATAAAAATCGTTCGTATAAATTTC	2364
RESULT 9			
ID	ADPS6127	standard; cDNA; 2744 BP.	
XX	AC	ADPS6127;	
XX	DT	18-NOV-2004 (first entry)	
DE		Human PRO CDNA sequence SEQ ID NO:2103.	
XX	KW	human; PRO; immune related disease; inflammatory immune response;	
KM		immune response stimulation; antiallergic; antianemic; antiarthritis;	
KM		antihistaminic; antidiabetic; antiinflammatory; antipsoriatic;	
KW		antiathermatic; antithyroid; CNS; dermatological; gastrointestinal;	
KW		haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;	
KM		nephroretropic; neuroprotective; osteopathic; respiratory; vasotrophic;	
KW		vitrucide; gene therapy; gene; ss.	
OS		Homo sapiens.	
XX	PN	WO2004039956-A2.	
XX	PD	13-MAY-2004.	
XX	PF	28-OCT-2003; 2003WO-US034381.	
PR		29-OCT-2002; 2002US-0422472P.	
PA	(GETH) GENENTECH INC.		
PI	Aggarwal S, Clark H, Gurney AL, Schoenfeld J, Williams PW;		
PT	Wood WJ, Wu TD;		
DR	MP1: 2004-376182/35.		
XX	P-P8DB; ADPS6128.		
XX	New PRO polynucleotides and polypeptides, useful in diagnosing		
PT	and treating an immune related disease, e.g. systemic lupus		
PT	erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in		
Pr	stimulating an immune response.		
XX	Claim 2; SEQ ID NO 2103; 3009pp; English.		
PS	The present invention describes an isolated PRO nucleic acid (1). Also		
XX	described: (1) a vector comprising (1); (2) a host cell comprising the		
CC	vector of (1); (3) a process for producing a PRO polypeptide; (4) an		
CC	isolated PRO polypeptide; (5) a chimeric molecule comprising the		
CC	polypeptide of (4) fused to a heterologous amino acid sequence; (6) an		
CC	antibody which specifically binds to a polypeptide of (4); (7) a		
CC	composition of matter comprising a polypeptide of (4), an agonist or		
CC	antagonist of the polypeptide or an antibody that binds to the		
CC	polypeptide in combination with a carrier; (8) an article of manufacture		
CC	comprising a container, a label on the container and a composition of		
CC	matter of (7); (9) a method of treating an immune related disease in a		
CC	mammal; (10) a method for determining the presence of a PRO polypeptide		
CC	in a sample suspected of having the polypeptide; (11) a method of		
CC	diagnosing an immune related disease or an inflammatory immune response		
CC	in mammal; (12) a method of identifying a compound that inhibits or		
CC	mimics the activity of or expression of a gene encoding a PRO polypeptide		
CC	; and (13) a method of stimulating the immune response in a mammal. The		
CC	PRO sequences have antiallergic, antianemic, antiarthritic,		
CC	antiaesthetic, antidiabetic, antiinflammatory, antipsoriatic,		

CC antirheumatic, antithyroid, CNS, dermatological, gastrointestinal, muscular,
 CC haemostatic, hepatotropic, immunostimulant, immunosuppressive, muscular,
 CC nephrotropic, neuroprotective, osteopathic, respiratory, vasotropic and
 CC virucide activities, and can be used in gene therapy. The nucleic acid
 CC (1) and the encoded polypeptides, compositions, kits and methods are
 CC useful in diagnosing and treating an immune related disease and in
 CC stimulating an immune response. The present sequence represents a human
 CC PRO nucleotide sequence from the present invention.

XX Sequence 2744 BP; 786 A; 585 C; 601 G; 772 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3,198-129	Length:	2744
Score:	1359.50	Matches:	301
Percent Similarity:	40.32%	Conservative:	1
Best Local Similarity:	40.19%	Mismatches:	0
Query Match:	85.02%	Indels:	447
DB:	13	Gaps:	1

US-10-659-004-104 (1-302) x ADP56127 (1-2744)

QY	1	MetaIaArgLyLeuSerValIleLeuIleLeuThrPheAlaLeuSerValThrAsnPro	20
DB	138	ATGGCAGAGAGTATCTGTATCTTGATCTGACCTTGGCCCTCTGTGCAAAATCC	197
QY	21	LeuHISGLuLeuValAlaAlaPheProGlnThrGluLysIleSerProAsnTrp	40
DB	198	CTTCATGAACTTAAAGACAGCTGCTTCCCGACAGACCTGAGAAAATGATCCGATTTGG	257
QY	41	GlusErGlyIleAsnValAspLeuAlaIleSerThrArgLysIleGlnGlnLeu	60
DB	258	GAATGTGGATTAATGTTGACCTTGGCAATTTCCACGCAATATCATCTCAACAGCTT	317
QY	61	PheTyraGlyGlyGluAsnAsnSerLeuSerValGluGlyPheArgLysLeuGln	80
DB	318	TTCCTACCGTATGAGAAAATATCTTGTGATGGAAGGCTTCAGAAAATTAATTCTCA	377
QY	81	AsnIleGlyIleAspLysIleLysArgIleLysIleIleHisIleAspHisAspHisSer	100
DB	378	AATATAGGATATGATTAAGTTAAAGAAATCCATATACCATGACACGACCATCATCTCA	437
QY	101	AspHisGluHisIleSerAspHisGluArgHisSerAspHisGluHisIleSerAsp---	119
DB	438	GACCAAGACATCATCTCAGACCATAGAGCTCCTCAGACCATGACATCATCTCAGACAC	497
QY	119	-----	119
DB	498	GAGCATCACTGACCATATATCATGCTGCTTGTGTAATAAATAGCAAAAGCTTTTGC	557
QY	119	-----	119
DB	558	CCAGACATGACTCAGATATGTTCAAGTAAAGATCTTAGAAACAGCCAGGGAAGAGCT	617
QY	119	-----	119
DB	618	CACGACCAAGAACATGCGAGTGTAGAAAGATGTCAGAGACAGTGTAGTAA	677
QY	119	-----	119
DB	678	GTGACCTCACTGTGTACAACTGTCTGTAGAGAACTCACTTTCTAGAGCAATAGAG	737
QY	119	-----	119
DB	738	ACTCCAAAGCTGAAAACTTTCCCAAGATGTAGAGAGCTCCATCCACCCAGTGTG	797
QY	119	-----	119
DB	798	ACATCAAGAGCCGGGTAGCCGCTGGCTGTAGAAAACAATGAATCTGTAGTGTAG	857
QY	119	-----	119
DB	858	CCCCGAAAAAGCTTTATGTATTCAGAAAAACAAATGAATAATCTCAGAGAGTGTTCAT	917

QY	119	-----	119
DB	918	GCATCAAGTACTGACATCTCATGATGGCATGCCATCCAGTTCCGCTGAATGCAACAGAG	977
QY	119	-----	119
DB	978	TTCAACTATCTCTGTCCAGCCATCATCAACCAATTGATGTAGATCTTGTCTAATTCAT	1037
QY	119	-----	119
DB	1038	ACAAGTAAAAAGAGCTGAATTCCTCCAAAGACTATTTCATTACAAATAGCCTGGGTT	1097
QY	119	-----	119
DB	1098	GGTGTATTATAGCATTTCCATCATCAGTTTCTGTCTGTCTGTGGGGTTATCTTAGTG	1157
QY	119	-----	119
DB	1158	CCTCTCATGAATCGGGTGTTTTCAATTTCTCTGAGTTTCTTGTGGACCTGGCGGT	1217
QY	119	-----	119
DB	1218	GGGACTTTGAGTGTGATGCTTTTACACCTTCTCCATTTCTCATGCAATGCACAC	1277
QY	119	-----	119
DB	1278	CATATCATAGCCATGAAGAACAGCAATGGAATGAAGAGAGCACTTTTCAGTCAT	1337
QY	119	-----	119
DB	1338	CTGTCTTCAAAACATAGAAAGAAAGTCCATTTTATTCACAGTGAAGGCTTAACA	1397
QY	119	-----	119
DB	1398	GCTTAGAGGCTGTATTTATGATTTCTTGTGACATGTCCTCATTTGATCAACAA	1457
QY	119	-----	119
DB	1458	TTTAAAGTAAAGAAAGAAAGATCAGAAAGAACTGAAATGATGATGTGAGATT	1517
QY	119	-----	119
DB	1518	AAGAAAGATTGTCAGATGATGATCTCACTTTCAACAAATGAGAGAAAGTAGATCA	1577
QY	119	-----	119
DB	1578	GATGATGAACGTAAAGCTATTACAGAGACTCAGAGAGCCCTCCACTTTGATTTCT	1637
QY	119	-----	119
DB	1638	CAGAGCTGCACTTGTGAAGAAAGAGATCATGATAGCTCATCTCATCCACAGAA	1697
QY	119	-----	119
DB	1698	GTCTACATGAATATGATCCAGAGGGTGCAGAAATTAATGCATTCACATTCACAGAT	1757
QY	119	-----	119
DB	1758	ACACTGGCAGTCAGACGATCTCATTCACACCATCATGATCAATATTTCTTCAT	1817
QY	120	-----HisIleProHisSerHisSerGlnArgLysArgGlu	1877
DB	1818	CATCAACACCAACCAACCATCTCTCAGATGACAGCCAGCTCTCTGGGAGAGAG	1937
QY	134	LeuLysAspAlaGlyValAlaThrLeuAlaTrpMetValIleMetGlyAspGlyLeuHis	1997
DB	1878	CTGAAGATGCGGCGTGCACATTTGGCTGTGATGTAAATGAGGTGATGCTGCAC	1957
QY	154	AsnPheSerArgGlyLeuAlaIleGlyAlaAlaPheThrGluGlyLeuSerSerGlyLeu	1917
DB	1938	AATTCAGCATGGCTTGAAGAAATTTGCTGCTTTTATCTGAAGGCTTATCAAGTGTTTA	1977
QY	174	SerThrSerValAlaValPheCysHisGlyLeuProHisGlyLeuGlyAspPheAlaVal	1937

Db 1998 AGTACTTCTGTTGCTGTGTGTGTATGATGTTGCTCAATGATTAAGTGAATTTGCTGT 2057
 Oy 194 LeuLeuValAGLMeThValValValValValValValValValValValValValVal 213
 Db 2058 CTACAAAGGCGGCGATACCGTTAAGCAGCGCTGCTTTAAAGCAATGTCACCGCATG 2117
 Oy 214 LeuAlaValLeuGLMeAlaValLeuValLeuValLeuValLeuValLeuValLeuVal 233
 Db 2118 CTGGGCTATCTTGGAAATGGCAACAGAAATTTTCATTTGCTATTCGTAATATGTTCT 2177
 Oy 234 MetTrpIlePheAlaLeuThzAlaValLeuPheMetHsValAlaLeuValAspMetVal 253
 Db 2178 ATGTGATATTTGCACTTACTGCTGCTTATTCATGTAATGTTGCTGCTGATATGTA 2237
 Oy 254 ProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSerArgTrpGlyTrpPhePhe 273
 Db 2238 CCTGAAAGCTGCGCAATGATGCTAGTGCATGATGATGATGATGATGATGATGATGAT 2297
 Oy 274 LeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeuLeuIleSerIlePheGlu 293
 Db 2298 TTACGAATATGCTGGGATGCTTTGGGTTTGGAAATTAATTAATTAATTAATTAATTA 2357
 Oy 294 HisValIleValPheArgIleAsnAsp 302
 Db 2358 CATAAATCGTCTTCTGTAATATTC 2384

RESULT 10
 ADR66497
 ID ADR66497 standard; DNA; 2744 BP.
 AC ADR66497;
 XX 02-DEC-2004 (first entry)
 DT
 XX Human prostatic carcinoma derived DNA SEQ ID 9 #3.
 DE
 XX human; cyrostatic; diagnosis; prostatic cancer;
 KM differential expression analysis; ds.
 OS Homo sapiens.
 OS
 PN MO2004076614-A2.
 XX 10-SEP-2004
 PD
 XX 22-FEB-2004; 2004MO-DE000433.
 PF
 XX 27-FEB-2003; 2003DE-01009985.
 PR 14-MAY-2003; 2003DE-01022134.
 RR
 XX (HINZ/) HINZMANN B.
 PA (DAHL/) DAHL E.
 PA (ROSE/) ROSENTHAL A.
 PA (HERM/) HERMANN K.
 PA (PILAR/) PILARSKY C.
 XX
 PI Hinzmann B, Dahl E, Rosenthal A, Hermann K, Pilarsky C, Specht T;
 PI Schmitt A, Beckmann G, Bruemendorf T, Kinneemann H, Roepcke S;
 PI Kinzhong L, Staub E,
 XX WPI; 2004-653386/63.
 DR
 XX New nucleic acids, and encoded proteins, from prostatic cancer tissue,
 PT useful for diagnosis, treatment and in screening for specific binding
 PT agents.
 XX
 PS Claim 1; Page 1130; 1607pp; German.
 CC This invention describes novel cyrostatic polynucleotide and polypeptide
 CC sequences which can be used in a method for diagnosing prostatic cancer
 CC or the risk of developing prostatic cancer. Diagnosis is based on
 CC determining over transcription or over expression of the sequences in

CC prostatic tissue. Screening for inhibitors of the sequences or detection
 CC substances involves a binding assay, any compounds that bind are
 CC selected, optionally after deconvolution of mixtures. Detection of a
 CC predetermined minimum level of the reporter indicates the presence of
 CC tumour cells. Inhibitors can be chosen from antisense oligonucleotides,
 CC short-interfering RNA or ribozymes; an organic molecule of molecular
 CC weight below 5000, preferably 300, that binds to the polypeptide; an
 CC aptamer against the polypeptide; a (monoclonal) antibody (Ab) against the
 CC polypeptide, preferably humanised or human; an anti-idiotypic, non-human
 CC (monoclonal) antibody directed against Ab or any of the above derivatised
 CC with a reporter group, cell toxin, immunostimulatory molecules and/or
 CC radioisotope. The polynucleotides are identified in human prostatic
 CC cancer by differential expression analysis, using DNA microarrays,
 CC between normal and tumorous tissues, with (over)expression being detected
 CC by quantitative PCR. Analysis of prostatic cancer samples showed that
 CC CD24 was upregulated in many of them. Sections of tissue, isolated from
 CC prostatic cancer patients, or subjects at risk, were incubated
 CC sequentially with anti-human CD4 murine monoclonal antibodies;
 CC biotinylated second antibody; streptavidin-conjugated horseradish
 CC peroxidase and then diaminobenzidine as colour former (brown). The
 CC samples were counterstained with hemalum (blue). Malignant cells stained
 CC strongly but non-malignant cells only weakly. In 15 of 63 samples of
 CC adenocarcinoma, membrane and cytoplasmic staining was very strong, and
 CC lymph node metastases were also stained. ADR65805-ADR66954 represent the
 CC polynucleotide and polypeptide sequences used in the method of the
 CC invention.

SQ Sequence 2744 BP; 786 A; 585 C; 601 G; 772 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 3,19e-129 Length: 2744
 Score: 1359.50 Matches: 301
 Percent Similarity: 40.32% Conservative: 1
 Best Local Similarity: 40.19% Mismatches: 0
 Query Match: 85.02% Indels: 447
 DB: 13 Gaps: 1

US-10-659-004-104 (1-302) x ADR66497 (1-2744)

Oy 1 MetAlaArgLysLeuSerValIleLeuThzPheAlaLeuSerValThrAsnPro 20
 Db 138 ATGGGAGGAAGTTATCTGTAATCTTGATCTGACCTTGGCCCTCTGTCAAAATGCC 197
 Oy 21 LeuHisGluLeuValAlaAlaPheProGlnThrThzGluValIleSerProAsnTrp 40
 Db 198 CTTCAATGAATCAAAAGCAGCTGCTTCCCGAGACCACTGAGAAATTAATGCGAATTGG 257
 Oy 41 GluSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTrpHisGluGlnLeu 60
 Db 258 GAATCTGGCAATTAATGTTGATGGCAATTTCCACAGGCAATATCAATCAACAGCTT 317
 Oy 61 PheTrpArgTrpGlyGluAsnAsnSerLeuSerValGluGlyPheArgLysLeuGln 80
 Db 318 TTCTACCGCTATGAGAAATAATTTCTTGCAATTAAGAGGTTCAAGAAATTAATTTCAA 377
 Oy 81 AsnIleGlyIleAspValIleLysArgIleHisIleHisAspHisAspHisSer 100
 Db 378 AATATAGGCAATAGATTAAGATTAAGATCAATTAACCATGACCAACCATCACTCA 437
 Oy 101 AspHisGluHisHisSerAspHisGluArgHisSerAspHisGluHisSerAsp 119
 Db 438 GACCACGACATCACTCAAGACATGACGCTCACTCAAGACATCACTCAAGACAC 497
 Oy 119 ----- 119
 Db 498 GAGCATCACTGACCAATATCATGCTGCTTCTGTAAATAAGCGAAAGCTCTTGGC 557
 Oy 119 ----- 119
 Db 558 CCAAGCATGACTCAGATAGTTTCAAGTAAAGATCTTAGAAAACAGCCAGGGAAGAGACT 617
 Oy 119 ----- 119

Db	618	CACCGACCAAGAACTGCGACGTGTAGAAAGAAATGTCAAGGACAGTGTAGTCTAGTAA	677
OY	119	-----	119
Db	678	GTGACCTCACTGTGTATCAACACACTGTCTGTAGAGAACTCACTTTCTAGAGACAATAG	737
OY	119	-----	119
Db	738	ACTCCAGACCTGGAAAACTTTCGCCAAGATGTAAAGCAGCTCCACTCCACCAGTGT	797
OY	119	-----	119
Db	798	ACATCAAAAGACCGGGGTAGCCGGGTGGCTGTGTAGAAAACAATGAATCTGTAGTGAG	857
OY	119	-----	119
Db	858	CCCCGAAAAAGCTTTATGTATTCAGAAACAATAATGAATCTCAGAGTGTTCAT	917
OY	119	-----	119
Db	918	GCATCAAAAGCTAGTGCATCTCATGGCATGGGCATCCAGTTCCGCTGAATGCACAGAG	977
OY	119	-----	119
Db	978	TTCAACTATCTGTGTCCAGCATCATCAACAATGTATGTCTAGATCTGTGTATTCAT	1037
OY	119	-----	119
Db	1038	ACAAGTAAAAAGAGGCTGAATCCCTCCAAAGACTTATCATTAACAATAGCCTGGTT	1097
OY	119	-----	119
Db	1098	GGTGTTTTAATGCCATTCCATCATCATCATTTCTGTCTCTGTGGGGTTATCTTAGTG	1157
OY	119	-----	119
Db	1158	CCTCATGAATCGGGGTGTTCCTCAATTTCTCGTAGTTCTTGTGGCAGTGGCGTT	1217
OY	119	-----	119
Db	1218	GGGACTTTAGTGTGTATGTCCTTTTACACCTTCTCCACATCTCATGCAAGTCAACAC	1277
OY	119	-----	119
Db	1278	CATAGTCATAGCCATGAAGACCAGCAATGGAATGAAAAAGACCACTTTTCACTCAT	1337
OY	119	-----	119
Db	1338	CTGTCTTCTCAAAACATGAAGAAGTGCTATTTTGAATTCGAGTGAAGGTCTAACA	1397
OY	119	-----	119
Db	1398	GCTTAGAGGCGCTGTATTTTCATGTTCTTGTGAACATGTCCTCATGTGATCAACAA	1457
OY	119	-----	119
Db	1458	TTTAAAGATAGAGAAAAAGATCAGAGAAACCTGAAAATGATGATGATGTGAGATT	1517
OY	119	-----	119
Db	1518	AAGAAAGACTGTCCAGATGATCTCACTTTCACAAATGAGAGAAAGTAGATACA	1577
OY	119	-----	119
Db	1578	GATGATCGAACTGAAGGCTATTTACGAGAGACTACAAAGACCTCCACTTTGATTTCT	1637
OY	119	-----	119
Db	1638	CAGCAGCTGCAGTCTTGAGAGAGAGAGATCATGATAGCTCATGTCAACAGAA	1697
OY	119	-----	119
Db	1698	GTCTTCATGAATATGTACCCGAGAGGGTGCAGAAATTAATGCATTCACATTTCCAGAT	1757

QY	119	-----	119
Db	1758	ACACTGGCCAGTCAGACGATCTCATTCACACCATCATGACTACATCATATTCTCCAT	1817
QY	120	-----HisHisProHisSerHisSerGlnArgTyrSerArgGluGlu	133
Db	1818	CATCACACACACCAAAACACCATCTCTACAGTCACAGCCAGCCGCTCTCTCGGGAGAGAG	1877
QY	134	LeuLysAspAlaGlyValAlaThrLeuAlaTrpMetValIleMetGlyAspGlyLeuHis	153
Db	1878	CTGAAGAATGCCGCGCTCGCCACTTGTGACCTGGATGGGATTAATGGGTGATGGCTTGAC	1937
QY	154	AsnHisSerAspGlyLeuAlaIleGlyValaAlaIleAspThrGluGlyLeuSerSerGlyLeu	173
Db	1938	AATTTCACGAGTAGGCCCTAGACATTTGGTGCTCTTTACTGAAGCTTATCAAGTGGTTTA	1997
QY	174	SerThrSerValAlaValPheCysHisGluLeuProHisGluLeuGlyAspPheAlaVal	193
Db	1998	AGTACTCTCTGTGCTGTGTTCTGTCTGATCAGATGTGCTCATGATTAATGAGTATGGATCTGTGTT	2057
QY	194	LeuLeuValAlaGlyMetThrValIleGlnAlaValLeuTyrAsnAlaLeuSerAlaMet	213
Db	2058	CTACTAAAGGCTGGGATGAGCCGTTTAAGCAGGCTGTCTTTAATGCAATTGTCAGCCATG	2117
QY	214	LeuAlaTyrLeuGlyMetValaThrGlyIlePheIleGlyHisTyrAlaGlyAsnValSer	233
Db	2118	CTGGGTCATCTTGGATATGGCAACAGAAATTTTCATTGGTCAATTATGCTGAATAATGTTCT	2177
QY	234	MetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisValAlaLeuValAspMetVal	253
Db	2178	ATGTGGATATTTGGACCTTACTGCTGGCTTATTCTATGATGTGCTGGTGGATATGGTA	2237
QY	254	ProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSerArgTyrGlyTyrPhePhe	273
Db	2238	CTGTAAATCTGCACATAGATGCTTAGTACCATGATGATACCGCTGGGGGATTTCTTT	2297
QY	274	LeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeuLeuIleSerIlePheGlu	293
Db	2298	TTACAGATATGCTGGGATGCTTTGGGTTTGGAAATTATGTTACTTATTTCCATATTGAA	2357
QY	294	HisLysIleValaPheArgIleAsnPhe	302
Db	2358	CATAAATCGTGTTCGTATAAATTTTC	2384
RESULT 11			
AD66155			
ID	AD66155	standard; DNA; 2744 BP.	
AC	AD66155;		
XX			
DT	02-DEC-2004	(first entry)	
XX			
DE	Human prostatic carcinoma derived DNA SEQ ID #2.		
XX	human; cytostatic; diagnosis; prostatic cancer;		
KW	differential expression analysis; ds.		
XX			
OS	Homo sapiens.		
XX			
PN	WO2004076614-A2.		
XX			
PD	10-SEP-2004.		
XX			
PF	22-FEB-2004; 2004WO-DE000433.		
XX			
PR	27-FEB-2003; 2003DE-01009985.		
PR	14-MAY-2003; 2003DE-01022134.		
XX			
PA	(HINZ/) HINZMANN B.		
PA	(DAHL/) DAHL E.		
PA	(ROSE/) ROSENTHAL A.		
PA	(HERM/) HERMANN K.		

PA (PILA/) PILARSKY C.
XX
PI Hitzmann B, Dahl E, Rosenthal A, Hermann K, Pilarczyk C, Specht T;
PI Schmitt A, Beckmann G, Bruemendorf T, Kinnemann H, Roepcke S;
PI Kinzhong L, Staub E;
XX
DR WPI; 2004-653386/63.
XX
PT New nucleic acid, and encoded proteins, from prostatic cancer tissue,
PT useful for diagnosis, treatment and in screening for specific binding
PT agents.
XX
PS Claim 1; Page 505; 1607pp; German.
XX
CC This invention describes novel cytoskeletal polynucleotide and polypeptide
CC sequences which can be used in a method for diagnosing prostatic cancer
CC or the risk of developing prostatic cancer. Diagnosis is based on
CC determining over transcription or over expression of the sequences in
CC prostatic tissue. Screening for inhibitors of the sequences or detection
CC substances involves a binding assay, any compounds that bind are
CC selected, optionally after deconvolution of mixtures. Detection of a
CC predetermined minimum level of the reporter indicates the presence of
CC tumour cells. Inhibitors can be chosen from antisense oligonucleotides,
CC short-interfering RNA or ribozymes; an organic molecule of molecular
CC weight below 5000, preferably 300, that binds to the polypeptide; an
CC aptamer against the polypeptide; a (monoclonal) antibody (Ab) against the
CC polypeptide, preferably humanised or human; an anti-idiotypic, non-human
CC (monoclonal) antibody directed against Ab or any of the above derivatised
CC with a reporter group, cell toxin, immunostimulatory molecules and/or
CC radioisotope. The polynucleotides are identified in human prostatic
CC cancer by differential expression analysis, using DNA microarrays,
CC between normal and tumour tissues, with (over)expression being detected
CC by quantitative PCR. Analysis of prostatic cancer samples showed that
CC CD24 was upregulated in many of them. Sections of tissue, isolated from
CC prostatic cancer patients, or subjects at risk, were incubated
CC sequentially with anti-human CD4 murine monoclonal antibodies;
CC biotinylated second antibody; streptavidin-conjugated horseradish
CC peroxidase and then diaminobenzidine as colour former (brown). The
CC samples were counterstained with hemalum (blue). Malignant cells stained
CC strongly but non-malignant cells only weakly. In 15 of 63 samples of
CC adenocarcinoma, membrane and cytoplasmic staining was very strong, and
CC lymph node metastases were also stained. AD65805-AD65854 represent the
CC polynucleotide and polypeptide sequences used in the method of the
CC invention.
XX
XX
SQ Sequence 2744 BP; 786 A; 585 C; 601 G; 772 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 3,19e-129 Length: 2744
Score: 1359.50 Matches: 301
Percent Similarity: 40.32% Conservative: 1
Best Local Similarity: 40.19% Mismatches: 0
Query Match: 85.02% Indels: 447
DB: 13 Gaps: 1
US-10-659-004-104 (1-302) x AD66155 (1-2744)
QY 1 Mch1arglyleuSerValileu1leuThrPheAlaLeuSerValThrAsnPro 20
DB 138 ATGCGAGAGAGATTATCTGTATCTGTATCTGACCTTGGCCCTCTGTCAAAATCCC 197
QY 21 LeuH1egluLeuValAlaAlaPheProGlnThrGluValylSerProAsnTrp 40
DB 198 CTTATGAACTTAAAGACGCTGCTTCCCAACACACTGAGAAATTAATGCGGATTGG 257
QY 41 GluSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTyrHisLeuGlnGlu 60
DB 258 GAATCTGCAATTAATGTTGACTTGGCAATTCACACGCAATATCACTCAACAGCTT 317
QY 61 PheTyrArgTyrGlyGluAsnAsnSerLeuSerValGluGlyPheArgLysLeuGln 80
DB 318 TTCTACCGCTAATGAGAAATATATTCTTGTCACTGAAGGTTCAAGAAATTAATTCAA 377

QY 81 AsnIleGlyIleAspLysIleLysArgTlleHisIleHisAspHisAspHisSer 100
DB 378 AATATAGCATATGATTAATTAATTAAGATTCATATACCCATGACCAACATCATCA 437
QY 101 AspHisGluHisHisSerAspHisGluArgHisSerAspHisGluHisHisSerAsp 119
DB 438 GACCAAGCATCTCACTGACATGAGCAATGAGGTCTACACCATGACATCACTCAGACCAC 497
QY 119 ----- 119
DB 498 GAGATCACTCTGACATATCATGCTGCTTCTGTGTAATAATAGCAAAAGCTTTGC 557
QY 119 ----- 119
DB 558 CCAAGCATGACTGATAGTATGTTCAAGTAAGATCTTGAAGAACAGCCAGGAAAGAGCT 617
QY 119 ----- 119
DB 618 CACCGACCAAGAACATGCCAGTGTGAAGAAATGTCAGACAGTGTAGTGTGA 677
QY 119 ----- 119
DB 678 GTGACCTCACTGTGTACAACTGTCTGTAGGAACCTCTTAGAGACAATAGAG 737
QY 119 ----- 119
DB 738 ACTCCAGACCTGGAAAACTTCTCCCAAGATGAAGACCTCCATCCACCCAGTGTG 797
QY 119 ----- 119
DB 798 ACATCAAAAGACCGGGTGAAGCCGCTGCTGTAGAAAAACAATGAATCTGTAGTGA 857
QY 119 ----- 119
DB 858 CCGGAAAAAGCTTATGTATTTCCAGAAAAACAATGAATTAATCTCAGAGTGTTCAT 917
QY 119 ----- 119
DB 918 GCATCAAAAGCTATGACATCTCATGCGATGGGATCCAGGTTCCGTGAATGCAACAGAG 977
QY 119 ----- 119
DB 978 TTCAACTATCTGTCTCAGCATCATCAACCAATTAATGCTAGATCTTGTCTGATTCAT 1037
QY 119 ----- 119
DB 1038 ACAAGTAAAGAGGCTGAATCCCTCCAAAGACTTATTCATTAACAATAGCCTGGTT 1097
QY 119 ----- 119
DB 1098 GGTGTTTATATAGCATTTTCATCATCATGATTTCTGTCTGTGGGGTTATCTTAGTG 1157
QY 119 ----- 119
DB 1158 CCTCTCATGAATCGGGTGTTCCTCAATTTCTCTGAGTTCTTTGCGACATGCGGCTT 1217
QY 119 ----- 119
DB 1218 GGAAGTTTGAAGTGAATGCTTTTACACCTTCTTCACATTCATGCAAGTCAACAC 1277
QY 119 ----- 119
DB 1278 CATATGATATGACATGAAGAACAGCAATGGAATGAAGAGAGCACTTTTCAATCAT 1337
QY 119 ----- 119
DB 1338 CTGCTTCTCAAAACATAGAAAGAAAGTCCATTTTGAATTCACAGTGAAGGCTTAACA 1397
QY 119 ----- 119
DB 1398 GCTCTAGAGGCTGTATTTTCTGTGTAATGTAATGCTCTCATGATGATCAACAA 1457
QY 119 ----- 119

OS Homo sapiens.
XX US2003099974-A1.
XX 29-MAY-2003.
XX 18-JUL-2002; 2002US-00198846.
XX 18-JUL-2001; 2001US-0306220P.
XX (MILL-) MILLENNIUM PHARM INC.
XX Lillie J, Xu Y, Wang Y, Steinmann K;
XX WPI, 2003-787014/74.
XX
XX Novel isolated polypeptide associated with breast cancer, useful for
XX detecting presence of polypeptide in sample, as a marker for breast
XX cancer.
XX
XX Disclosure; SEQ ID NO 10007; 36pp; English.
XX
XX The invention relates to an isolated polypeptide (I) associated with
XX breast cancer which is encoded by a nucleic acid molecule comprising a
XX nucleotide sequence (SI). Further disclosed is an antibody that binds to
XX the polypeptide of the invention. The activity of the polypeptide of the
XX invention may be described as cytostatic. The antibody is useful for
XX detecting the presence of (I) in a sample. Nucleic acid molecules of the
XX invention are useful in the detection of breast tumors. (I) is useful as
XX a marker for breast cancer and in breast cancer therapy. Sequences given
XX in records ACN78851-ACN92934 represent nucleic acid markers associated
XX with breast cancer. Note: The sequence listing does not form part of the
XX specification but may be obtained in electronic format from the USPTO web
XX site at seqdata.uspto.gov/sequence.html?docid=20030099974
XX
SQ Sequence 3537 BP; 1060 A; 707 C; 747 G; 1016 T; 0 U; 7 Other;
XX
Alignment Scores:
Pred. No.: 4,516-129 Length: 3537
Score: 1359.50 Matches: 301
Percent Similarity: 40.32% Conservative: 1
Best Local Similarity: 40.19% Mismatches: 0
Query Match: 85.02% Indels: 447
DB: 11 Gaps: 1
US-10-659-004-104 (1-302) x ACN88857 (1-3537)
QY 1 MetAlaArgLysLeuSerValIleLeuIleLeuThrPheAlaLeuSerValThrAspPro 20
DB 138 ATGGCGAGGAGTATTCTGTATCTTGATCTGACCTTGCCCTCTGTGTCACAAATCC 197
QY 21 LeuHISGLIuLeuLysAlaAlaAlaPheProGlnThrGluLysIleSerProAsnTrp 40
DB 198 CTTTCATGAACTAAAGCAGCTCTTCCCGCAGACCATGAGAAATTAAGTCCGATGG 257
QY 41 GluSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTyrHisIleGlnGlnLeu 60
DB 258 GAATCTGGCATTAATCTTGCATCTTGGCAATTTCCACCGGCATTAATCACTACAAAGCTT 317
QY 61 PheTyrArgTyrGlyGluAsnAsnSerLeuSerValGlnGlyPheArgLysLeuLeuGln 80
DB 318 TTCCTACCGCTATGAGAAATAATCTTGTGTCAGTTGAAGGGTTCAGAAATTAATCTTCA 377
QY 81 AsnIleGlyIleAspLysIleLeuArgGlnIleHisIleHisAspHisAspHisSer 100
DB 378 AATATAAGGCGATGATTAAGATTAAGAAATCCATATACACCATGACACGACCATATCA 437
QY 101 AspHisGluHisHisSerAspHisGluArgHisSerAspHisGluHisSerAsp--- 119
DB 438 GACCGAGGAGCATCTCAGACCATGAGCTCTCAGACCATGAGCATCACTCAGACCAAC 497
QY 119 ----- 119

DB 498 GAGCATCACTGTGACCATTAATCATGTCTTCTGTGTAATAAATAGCGAAAGCTTTTGC 557
QY 119 ----- 119
DB 558 CCAGACCATGACTGAGATAGTTTCAGTAAAGATCTTAGAAACAGCAGGGGAAAGAGCT 617
QY 119 ----- 119
DB 618 CACCGACCAAGACATGCGAGTGTGAAGAAATGTCAAGACAGTGTAGTCTTAGTGAA 677
QY 119 ----- 119
DB 678 GTGACTCACTGTGTACAACACTGTCTGTGAAGAACTCACTTTAGAGACAATAGAG 737
QY 119 ----- 119
DB 738 ACTCCAGACCTGGAAAACTTCCCAAGATGTAGACGCTCCACTCCAGAGTGTG 797
QY 119 ----- 119
DB 798 ACATCAAGAGCCGGGTGAGCCGGCTGGCTGTAGAAAACAAATGAATCTGTGATGAG 857
QY 119 ----- 119
DB 858 CCGGAAAGGCTTTATGTAATTCAGAAACAAATGMAATCTCAGAGTGTTCAT 917
QY 119 ----- 119
DB 918 GCATCAAGACTACTGACATCTCATGAGATGGGATCCAGATTCCGCTGAATGCAACAGAG 977
QY 119 ----- 119
DB 978 TTCAACTATCTGTCTCCAGCATATCATCAACCAATTGATGTATCTGTCTGATTCAT 1037
QY 119 ----- 119
DB 1038 ACAAGTAAAGAGGCTGAATCCTCCAAAGACTATTTCATTAACAATAGCTGGGT 1097
QY 119 ----- 119
DB 1098 GGTGTTTATAGCCATTTCCATCATGATGTTCTGTCTGTGCGGGTATCTTAGTG 1157
QY 119 ----- 119
DB 1158 CTTCTCATGAATCGGGTGTTCATTAATTTCTCTGAGTTTCTTGGCAGTGGCGGT 1217
QY 119 ----- 119
DB 1218 GGGACTTTGAGTGTATGCTTTTACACCTTCTCCACATTTCTCATGCAAGTCAACCAC 1277
QY 119 ----- 119
DB 1278 CATATCATAGCCATGAGAAACGACCAATGAAATGAAGAGGACCACTTTTCAATCAT 1337
QY 119 ----- 119
DB 1338 CTGTCTTCTCAAAATAGAAAGAAAGTCCATTTTGTGATTTCCAGTGGAAAGGTCTAACA 1397
QY 119 ----- 119
DB 1398 GCTCTAGAGGCTGTATTTTCATGTTCTTGTGMAATGTCCTCAGATTGATCAACAA 1457
QY 119 ----- 119
DB 1458 TTTAAGATTAAGAAAGAAAGAAATCAGAGAAACCTGAAATATGATGATGTGAGATT 1517
QY 119 ----- 119
DB 1518 AAGAAGCAGTTGTCCAAATGATGATTCACACTTTCAACAAATGAGAGAAAGTATATCA 1577
QY 119 ----- 119
DB 1578 GATGATCAACTGAAAGCTATTATTAAGAGACAGATCACAAGAGCCCTCCACTTTGATTC 1637

Dp	361	GAGCATCTGACCAATGATCATCTCCACCATTAATCATGTGCTTGGTAAAT	420
Qy	119	-----	119
Dp	421	AAGGAAAAGCTCTTTGCCAGACCATGACTCAGATGAGTAAAGATCTAAGAAC	480
Qy	119	-----	119
Dp	481	AGCCAGGGGAAAGAGCTCACCGACCAGAACATGCCAGTGGTAAGAAGAAATGCAAGAC	540
Qy	119	-----	119
Dp	541	AGTGTAGTCTAGTGAAGTGAACCTCAACTGTATCAACAACACTGTCTGAAGAACTCAC	600
Qy	119	-----	119
Dp	601	TTTCTAGAGACAATAGAGACTCCAAAGACTGGAAAACTTTCCCAAGATGTAAGCAGC	660
Qy	119	-----	119
Dp	661	TCCACTCCACCAGGTGCACATCAAAAGACCGGGTGAACCGGCTGGTGTAGAAAACA	720
Qy	119	-----	119
Dp	721	AATGATCTGTGAGTGAGCCCCGAAAAGGCTTTATGTATTCAGAAACAATAAGTAAAT	780
Qy	119	-----	119
Dp	781	CCTCAGAGTGTTCATATGATCAAAAGCTACTGATCATGGCATGGGATCCAGGTT	840
Qy	119	-----	119
Dp	841	CCGCTGAATGCAACAGAGTTCAACTATCTGTGCCAGCATATCAACCAATTGATGCT	900
Qy	119	-----	119
Dp	901	AGATCTTGCTGATTCATACAAGTGAAGAAGGCTGAATCCTCCAAAAGCTATTCA	960
Qy	119	-----	119
Dp	961	TTACAAATAGCCTGGGTGGTGGTTTAAAGCAATTCATATCATGATTCCTGTCTCG	1020
Qy	119	-----	119
Dp	1021	CTGGGGGTATCTTAGTGCCTCTCATGAAAGGGGTGTTTCAAAATTTCTCTGAGTTTC	1080
Qy	119	-----	119
Dp	1081	CTTGTCGACTGGCCGTTGGGACTTTGAGTGGTATGCTTTTACACTTCTTCCACAT	1140
Qy	119	-----	119
Dp	1141	TCTCATGCAAGTCACCAACCATATGTCATAGCCATGAAGAACGACATGAAATGAAGA	1200
Qy	119	-----	119
Dp	1201	GGACCACTTTCAATCATCTGTCTTCTCAAAACATAGAAGAAAGTGCATTTTGATTCC	1260
Qy	119	-----	119
Dp	1261	ACGTGAAGGCTTAACAGCTTAGAGGCTGTATTTCAATGTTTCTTGTGAACATGTC	1320
Qy	119	-----	119
Dp	1321	CTCATTTGATCAACAATTTAAAGATAAAGAAAAAAGATCAGAGAAACCTGAAAT	1380
Qy	119	-----	119
Dp	1381	GATGATGATGTGAGATTAAAGACAGTTGTCCAAATGAAATCTCAACTTTCACAAAT	1440
Qy	119	-----	119
Dp	1441	GAGGAAAGTAGATACAGATGATCGAACTGAAGGCTATTTCAGGCACTCACAGAG	1500

QY	119	-----	119
Db	1501	CCCTCCACATTGATTTCTACGACGCTGCAGTCTTGGAGAGAGAGATCATGATAGCT	1560
QY	119	-----	119
Db	1561	CATGCTCATCCACAGGAAGTCTCAATGAAATATGTACCCAGAGGTGCAGAAATTAATGC	1620
QY	119	-----	119
Db	1621	CATTCACTTTCCACGATACACTCGGCCAGTCCAGAGATCTCATTCACCACTCATGAC	1680
QY	120	-----	127
Db	1681	TACCATATATTTCTCCATCTATCCACACACACCAACCAACCATCTCTACAGCTCAGACCCAG	1740
QY	128	ATGTGTserArggIugIuleuLyAspAlaGlyValAlaThrLeuAlaIrrpMetValIle	147
Db	1741	CGGTACTCTCGGGAGGAGTGTGAAGAATGCCGGGTGCCACTTTGGCTGTGATGTGATG	1800
QY	148	MetGlyAspGlyLeuHlaEnPheSerAspGlyLeuAlaIleGlyValAlaIrrpMetArgIu	167
Db	1801	ATGGGTGATGGCTCTGCACAAATTCACGACGAGGCTTCGACAAATGGTCTGCTTTTACTGAA	1860
QY	168	GlyLeuSerSerGlyLeuSerThrSerValAlaValPheCysHisGlyLeuProHisGly	187
Db	1861	GGCTTATCAAGTGTGTTAAGTACTTCTGTTCGTGTCTGTCTGTATAGTTCCTCATGAA	1920
QY	188	LeuGlyAspPheAlaValLeuLeuLyAspGlyMetThrValIleGlyAlaValLeuTy	207
Db	1921	TTAGTGATCACTTGTCTGTTCTACTAAAGGCTGCATGACCGTTAAGCAGGCTGTCTTTAT	1980
QY	208	AspAlaLeuSerLamMetLeuAlaGlyLeuGlyMetAlaThrGlyIlePheIleGlyHis	227
Db	1981	AATGCAATTTGCAGCAGCATGCTGGGGATCTTGGAAATGCAACAGAAATTTTCATTGGTCAT	2040
QY	228	TyrAlaGlyAsnValSerMetCPLIlePheAlaLeuThrAlaGlyLeuPheMetHisVal	247
Db	2041	TATGCTGAATAATGTTCTCTATGTGGATATTTTGGCACTTACGCTGGCTTATTACATGATAGTT	2100
QY	248	AlaLeuValAspMetValProGlyMetLeuHisAsnAspAlaSerAspHisGlyCysSer	267
Db	2101	GCTCTGTGATATATGTGACCTGGAAGCTGCACAAATGATGCTGTGATACCACTTGATGTAGC	2160
QY	268	ArgTrrGlyTyrPhePheLeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeu	287
Db	2161	CGCTGGGGGATATTTCTTTTACGAATGCTGGGATGCTTTTGGGTTTGGAAATTAATGTTA	2220
QY	288	LeuIleSerIlePheGlyHisLyAspAlaGlyPheArgIleAsnPhe	302
Db	2221	CTTATTTCCATATTTGAAATTAATAATCGTGTCTTCGATTAATTAATTC	2265
RESULT 15			
ID	ADL57160	standard; DNA; 2268 BP.	
AC	ADL57160;		
XX	03-JUN-2004	(first entry)	
DT			
XX			
DE	Human NOV9d gene SEQ ID NO:105.		
XX			
KM	ds; gene; human; antidiabetic; anorectic; virinucide; cardiatic; hypotensive;		
KM	antiarterioesclerotic; anorectic; neuroprotective; antibacterial; fungicide;		
KM	prozoacide; nootropic; neuroprotective; antiparkinsonian;		
KM	anticonvulsant; osteopathic; antidiabetic; antiinflammatory;		
KM	dermatological; antiaesthetic; antidiabetic; gene therapy;		
KM	fibroblast growth factor receptor 4; FGFR4;		
KM	complement factor I precursor; matrix metalloproteinase-15 precursor;		
KM	MDG3; T-lymphocyte surface antigen Ly-9 precursor;		
KM	fibroblast growth factor-21; FGF-21;		
KM	alpha-2-macroglobulin-like polypeptide variant;		

KW antileukoprotease 1 precursor; LIV-1; nuclear hormone receptor NOR-1;
 KW transmembrane protein-like; beta-neoendorphin-dynorphin precursor.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 FT CDS 1..2268
 FT /tag= a
 XX MO2004022723-A2.
 XX 18-MAR-2004.
 PD 09-SEP-2003; 2003WO-US028141.
 PF 09-SEP-2002; 2002US-0409145P.
 PR 10-SEP-2002; 2002US-0409544P.
 PR 12-SEP-2002; 2002US-0410320P.
 PR 16-SEP-2002; 2002US-0411060P.
 PR 23-SEP-2002; 2002US-0412766P.
 PR 23-SEP-2002; 2002US-0412825P.
 PR 24-SEP-2002; 2002US-0412767P.
 PR 25-SEP-2002; 2002US-0413342P.
 PR 30-SEP-2002; 2002US-0414832P.
 XX (CURA-) CURAGEN CORP.
 PA Zhong M, Guo X, Anderson DW, Ort T, Padigaru M, Rieger DK;
 P1 WPI; 2004-315567/29.
 XX P-PSDB; ADL57161.
 DR New isolated NOVX polypeptides and polynucleotides, useful for
 FT preventing, diagnosing or treating NOVX-associated disorders, e.g.
 PT osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
 PT asthma, or infections.
 XX Claim 17; SEQ ID NO 105; 214pp; English.
 PS The invention relates to a novel isolated polypeptide (NOVX) comprising a
 CC mature form of any of the 37 amino acid sequences fully defined in the
 CC specification. A polypeptide of the invention has antidiabetic,
 CC anorectic, cardiant, hypotensive, antiatherosclerotic, anorectic,
 CC vitruce, antibacterial, fungicide, protozoacide, nootropic,
 CC neuroprotective, antiparkinsonian, anticonvulsant, osteopathic,
 CC antiarthritic, antiinflammatory, dermatological, antiasthmatic, and
 CC antipneumatic activity. A polynucleotide of the invention may have a use
 CC in gene therapy. The polypeptides, nucleic acid molecules and antibodies
 CC are useful in the manufacture of a medicament for treating a syndrome
 CC associated with a human disease, preferably a NOVX-associated disorder.
 CC The nucleic acid molecules, polypeptides and antibodies are useful for
 CC treating, preventing or diagnosing diseases such as metabolic disorders,
 CC diabetes, obesity, infectious diseases (viral, bacterial, fungal,
 CC helminthic, and protozoal), anorexia, cancer, cardiovascular diseases
 CC (hypertension, atherosclerosis), neurodegenerative disorders, Alzheimer's
 CC disease, Parkinson's disease, epilepsy, immune disorders
 CC (osteoarthritis), haematopoietic disorders, inflammatory skin disorders,
 CC asthma, and various dyslipidaemias. The nucleic acids and polypeptides
 CC may also be used as targets for the identification of small molecules
 CC that modulate or inhibit e.g. neurogenesis, cell differentiation, cell
 CC proliferation, haematopoiesis, wound healing and angiogenesis, in gene
 CC therapy, in generation of antibodies that bind immunospecifically to NOVX.
 CC Substances for use in therapeutic or diagnostic methods. The nucleic
 CC acids are further used as hybridisation probes, in chromosome mapping,
 CC tissue typing, preventive medicine, and pharmacogenomics. The NOVX
 CC polypeptides of the invention show homology to certain known human
 CC proteins: NOV9a-1c show homology to fibroblast growth factor receptor 4
 CC (FGFR4); NOV9a shows homology to complement factor 1 precursor; NOV3a
 CC shows homology to matrix metalloproteinase-15 precursor; NOV4a shows
 CC homology to MDC3; NOV5a-5c show homology to T-lymphocyte surface antigen
 CC Ly-9 precursor; NOV6a-6m show homology to fibroblast growth factor-21
 CC (FGF-21); NOV7a-7c show homology to alpha-2 macroglobulin-like
 CC polypeptide variant; NOV8a-8g show homology to antileukoprotease 1

CC precursor; NOV9a-9i show homology to LIV-1 protein; NOV10a shows homology
 CC to nuclear hormone receptor NOR-1; NOV11a-11j show homology to
 CC transmembrane protein-like; NOV12a-12c show homology to beta-neoendorphin
 CC -dynorphin precursor. The present sequence encodes a NOVX polypeptide of
 CC the invention.
 XX SQ Sequence 2268 BP; 661 A; 512 C; 476 G; 619 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 5,01e-129 Length: 2268
 Score: 1356.50 Matches: 301
 Percent Similarity: 40.00% Conservative: 1
 Best Local Similarity: 39.87% Mismatches: 0
 Query Match: 84.83% Indels: 453
 DB: 12 Gaps: 1
 US-10-659-004-104 (1-302) x ADL57160 (1-2268)
 QY 1 MetAlaArgLysLeuSerValIleLeuIleLeuThrPheAlaLeuSerValThrAsnPro 20
 DB 1 ATGGGAGAGAAATTATCTGTAATCTTGATCTTGACCTTGGCTTCTGTGACAAACCC 60
 QY 21 LeuHISGluLeuLysAlaAlaAlaPheProGlnThrTrpGluLysIleSerProAsnTrp 40
 DB 61 CTTGATGAACTTAAAGACAGCTGCTTCCCGACAGCCATGAGAAATTAGTCGAAATTGG 120
 QY 41 GluSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTrpHisLeuGlnLeu 60
 DB 121 GAATCTGGCATTAATGTTGATCTTGCAATTTCCACAGCGCAATATCATCTCAACAGCTT 180
 QY 61 PheTrpArgTrpGlyGluAsnAsnSerLeuSerValGluGlyPheArgLysLeuGln 80
 DB 181 TTCTACCGCTATGGAGAAATTAATCTTGTGAGTGAGGGGATTCAGAAATTAATTCTCA 240
 QY 81 AsnIleGlyIleAspLysValIleLysArgIleHisIleHisAspHisAspHisSer 100
 DB 241 AATATAGGCATATGATTAATTAATAAGATCATATATACACGACGACCATCATCA 300
 QY 101 AspHisGluHisHisSerAspHisGluArgHisSerAspHisGluHisHisSerAsp 119
 DB 301 GACCACGACATCACTCAGACCAATGAGGCTCATCAGACATGAGCATCACTCAGACCA 360
 QY 119 ----- 119
 DB 361 GAGCATCTCTGACCATGATCATCATCTTCACCAATATCATGCTCTTGTGTAATAAT 420
 QY 119 ----- 119
 DB 421 AAGGAAAAGCTCTTTGCCAGACCATGACTCAATAGTTCAGGTAAGATCTTAGAAAC 480
 QY 119 ----- 119
 DB 481 AGCCAGGGGAAAGAGCTCACCGACCAAGACATGCGAGGTGTAGAGAAATGTCAAGAC 540
 QY 119 ----- 119
 DB 541 AGTGTAAGTGTAGTGAAGTGAAGTCACTGTAACAACTGTCTCTGAAGAACTTCAC 600
 QY 119 ----- 119
 DB 601 TTTCTAGAGACATAGAGACTCCAGACCTGGAAAACTTTCCCAAGATGTAAAGCAGC 660
 QY 119 ----- 119
 DB 661 TCACATCCACCAAGTGTACATCAAAAGACCGGTGAGCCGGTGTGTAGGAAAACA 720
 QY 119 ----- 119
 DB 721 AATGAATCTGTAGTGAAGCCCGAAAGGCTTTATGTATTCAGAAAACAAATGAAT 780
 QY 119 ----- 119
 DB 781 CCTGAGAGTGTTCATGATCAAAAGTACTGATCATCTCATGCGATGGGATTCAGGTT 840

```

QY 119 ----- 119
Db 841 CCGGTGAATGCAAGAGTTCACTATCTGTCCAGCCATCATCAACCAATTGANGCT 900
QY 119 ----- 119
Db 901 AGATCTTGCTGATTCAATACAGTGAAGAAAGGCTGAATCCCTCCAAAGACTATTCA 960
QY 119 ----- 119
Db 961 TTACAAATAGCCTGGGTGGTGGTTTATAGCAATTCATCATCATGTTCTGTCTG 1020
QY 119 ----- 119
Db 1021 CTGGGGGTTATCTTAGTCCCTCTCATGAATGCGGTGTTTTCMAATTTCTCCTGAGTTTC 1080
QY 119 ----- 119
Db 1081 CTTGTGGCACTGGCCGTTGGGACTTTGAGTGTGATGCTTTTACACCTTCTTCCACAT 1140
QY 119 ----- 119
Db 1141 TCTCATGCAAGTCAACCACTATGATAGCCATGAAGAACCAAGCAATGAATAAGAA 1200
QY 119 ----- 119
Db 1201 GGACCACTTTTATGATCATCTGTCTTCTCAAAACATAGAAGAAAGCTTATTGATTCC 1260
QY 119 ----- 119
Db 1261 ACGTGAAGGGTCTAACAGCTCTAGAGGCTGTATTTTCATGTCTTCTTGTGAACATGTC 1320
QY 119 ----- 119
Db 1321 CTCACATTGATCAACCAATTTAAAGTAAAGAAAGAAAGATCAGAAAGAACTGAAAT 1380
QY 119 ----- 119
Db 1381 GATGATGATGTGAGATTAAAGACAGTTGTCCAAATGAAATCTCACTTTCAACAAAT 1440
QY 119 ----- 119
Db 1441 GAGGAGAAAGTATGATACAGATGCACTGAAGGCTATTTCAGAGAGACTCACAAAG 1500
QY 119 ----- 119
Db 1501 CCCTCCACTTGTATCTCAGCAGCTGAGTCTTGAAGAAAGAGGTCAATGATGCT 1560
QY 119 ----- 119
Db 1561 CATGCTCATCAAGAGAGTCAATGAATATGTACCAAGAGGTGCAAGAAATGAATGC 1620
QY 119 ----- 119
Db 1621 CATTCATTTCCAGATACACTCGGCGAGTCAAGAGATCTCATCAACCATCATGAC 1680
QY 120 ----- 120
Db 1681 TACCATCATATTTCTCATCATCAACCAACCAAAACCAACATCTCAGAGTCAAGCCAG 1740
QY 128 ArgTyrSerArgIuLuleuLysAspAlaGlyValAlaThrLeuAlaTrpMetValIle 147
Db 1741 CGCTACTCTCGGAGAGAGTGAAGAGTCCGGGCTGCGCACTTGCGCTGGAGTGGTATA 1800
QY 148 MetGlyAspGlyLeuHisAsenPheSerAspGlyLeuAlaIleGlyValAlaPheThrGln 167
Db 1801 ATGGGTGATGCTGCAATTTCAAGGATGGCTTAGCAATTGCTGCTGTTTACTGAA 1860
QY 168 GlyLeuSerSerGlyLeuSerThrSerValAlaValPheCysHisGlyLeuProHisGly 187
Db 1861 GGCTTATCAAGGTGTTAAGTACTTCTGTTGCTGTGTTCTGTCATGATGAGTTCCTCATGAA 1920

```

```

QY 188 LeuGlyAspPheAlaValLeuLeuLysAlaGlyMetThrValIleGlnAlaValLeuTyr 207
Db 1921 TTAGGTGACTTTGCTGTTCTACTTAAAGCTGGCATGACCTTAAGCAGGCTGTCCTTAT 1980
QY 208 AsnAlaLeuSerAlaMetLeuAlaTyrLeuGlyMetAlaThrGlyIlePheIleGlyHis 227
Db 1981 AATGCATTGTCAAGCATGCTGGCGTATCTTGAATGGCAACAGAAATTTTCATTGGTCAAT 2040
QY 228 TyrAlaGlnAsnValSerMetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisVal 247
Db 2041 TATGCTGAATAATGTTCTATGTGATATTTGCACTTACTGCTGCTTATTCATGATGTT 2100
QY 248 AlaLeuValAspMetValProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSer 267
Db 2101 GCTTGCTGATATGATGTAACCTGAATGCTGCACATGATGCTAGTGAACATGATGATG 2160
QY 268 ArgTyrGlyTyrPhePheLeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeu 287
Db 2161 CGCTGGGGGTATTTCTTTTACAGAAATGCTGGATGCTTTTGGGTTTGGAAATATGTTA 2220
QY 288 LeuIleSerIlePheGluHisIleValPheArgIleAsnPhe 302
Db 2221 CTATTTTCATATTTGAACATAAATGCTGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2265

```

Search completed: March 19, 2005, 23:11:38
 Job time : 644 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 19, 2005, 22:48:51, Search time 3692 Seconds
(without alignments)
3113.600 Million cell updates/sec

Title: US-10-659-004-104

Perfect score: 1599

Sequence: 1 MARKSLVLLTFLALSVTNP.....FGIMLLISFEHKVFRINF 302

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-Q/cgnt_1/USPro.epool/US10659004/unatc 14032005 181235 15755/app query.fasta_1.455
-DB=EST -QMT=fastap -SUFFIX=ref -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10659004.OCGN_1_1_3437 @tunatc 14032005 181235 15755 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

EST:
1: gb_est1.*
2: gb_est2.*
3: gb_hcc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gse81.*
9: gb_gse82.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1237.5	77.4	3834 3	AK028976 Mus muscu
2	1007	63.0	859 3	CA488910 AGENCOURT
3	981	61.4	864 6	CD251026 AGENCOURT
4	975	61.0	756 6	CA327621 UT-M-FRO-
5	974	60.9	906 6	BU541953 AGENCOURT
6	969	60.6	605 4	B1284823 UI-R-CX08
7	948	59.3	665 7	CF909638 A0539C05-
8	935	58.5	875 5	BX327149 BX327149
9	906	56.7	893 6	CA454784 AGENCOURT

10	893	55.8	582 5	BP259977
11	880	55.0	548 5	BU479906
12	875	54.7	791 4	BG547794
13	850	53.2	539 6	CB052402
14	850	53.2	639 5	BU614923 UI-M-FRO-
15	846	52.9	743 1	A1907176 RC-BT133-
16	838	52.4	746 4	BG530601 602560191
17	836.5	52.3	571 2	BB373870 601227885
18	812.5	50.8	1184 3	CR647282 Tetraodon
19	812	50.8	1216 3	CR646555 Tetraodon
20	806	50.4	1049 3	CR639910 Tetraodon
21	803.5	50.3	1134 3	CR636655 Tetraodon
22	801	50.1	857 2	BF032013 601559465
23	777	48.6	1059 4	BM557996 AGENCOURT
24	772	48.3	692 7	CF147123 UI-HF-CBO
25	771	48.2	691 7	CF146933 UI-HF-CBO
26	759.5	47.5	1135 4	BM560789 AGENCOURT
27	753	47.1	1134 4	BG994120 PMO-HT091
28	740	46.3	950 5	BQ433766 AGENCOURT
29	739	46.2	699 7	CN641564 ILLUMIGEN
30	738	46.2	486 2	BP913546 MR3-UT012
31	736	46.0	855 5	BU159290 AGENCOURT
32	728	45.5	849 2	BF680209 602154961
33	719	45.0	766 5	BM948026 UI-M-EGOP
34	717.5	44.9	640 7	CO354780 DR. ATE. FL
35	717	44.8	757 6	CD298707 AGENCOURT
36	713	44.6	5235 3	CR749813 Homo sapi
37	712	44.5	3706 3	AK049099 Mus muscu
38	711	44.4	480 2	BF913542 MR3-UT012
39	710	44.4	730 5	BX926167 BX926167
40	707	44.2	3924 3	AK030685 Mus muscu
41	701	43.8	580 2	BP915495 MR3-UT012
42	696	43.5	536 2	BF915454 MR3-UT012
43	695.5	43.5	751 7	CN526184 UI-M-HNO-
44	695.5	43.5	872 1	AU120027 AU120027
45	688	43.0	632 5	BX925865 BX925865

ALIGNMENTS

RESULT 1	AK028976	3834 bp	mRNA	linear	HTC 03-APR-2004
LOCUS	AK028976				
DEFINITION	Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched library, clone:473247B07 product:ESTROGEN REGULATED LIV-1 PROTEIN homolog [Homo sapiens], full insert sequence.				
ACCESSION	AK028976				
VERSION	AK028976.1	GI:26324937			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	Carninci, P. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	2				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				
REFERENCE	3				
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M., Smi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kaishogi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M.,				

TITLE
JOURNAL MEDLINE
20530913
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3834)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hoshizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Horii, F., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toy, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in RIKEN Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
FEATURES
SOURCE
1. 3834
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM:DB:4732477B07"
/db_xref="taxon:10090"
/clone="4732477B07"
/tissue_type="skin"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days neonate"
500..2797
/note="unnamed protein product; ESTROGEN REGULATED LIV-1 PROTEIN homolog [Homo sapiens] (SPR1013433, evidence: FASTA, 86.3%id, 100%length, match=2295)
putative
/codon_start=1
/protein_id="BAC26223.1"
/db_xref="GI:26324938"
/translation="MATDLSVIMLTPLAWTSPILHSTQTAASFQTEKINSWEPG VAVDAVTMORHLOQLPRYGENDSISVEGRKLQNIQIDKIRYVHHDEHSHAD HEHSDHEHSDHEHSDHEHSDHEHSDHEHSHRSHVTAQKNRRAFCDLSDSDN SGNKRTSLGKSRPAENHNRNRTKEASSESVTSYAVNAVSEGRVETLEPKPG RRTKDVNSPTSPITEKSRVGLSRILARKSNSESVSEPKRSRRNNNDIQECFNT TLTLSHGMSIALNATEFNYLCALINQIDARACLHTSEKKAELIPKTVSLQIA MLGGRFIAISIIISFLSLGVIIVPLMNRVPEFKRLSFLVALAVGTLSGDLHLPHSH

ASHOHSHSEEPAMKRGPLPSHLASNONIESSYPDSTWKGTLALGLYEMFLVHY
LTIKOPKDDKKKNOCKKRENDREVESKQLSKYDSQSSNEBKVDPEERPESTYRARS
OEPSFPDSQOPMLSEBEVITAAHPVEVNEVYPRCKRKHSHFDITGQSDLIH
HHHDYHHIILHHHROHHPHSHQVRSREELKQAGITLLAMVYMGGLNHSGLAI
GAFTBELSGLSVAVFCHELPHEIGDPAVLKAMTVKQAVLYNALSMALYLGM
ATGIFIGHYENVMNIFALTAGLFMYVALVDVPEMLHNDASDHGSRMGYFELQNA
GILLGFMILLISFEKIVIRINF"
3819..3824
/note="putative"
3834
polya_site
polya_signal
3834
/note="putative"
3834
ORIGIN
polya_site
/note="putative"
3834
Alignment Scores:
Pred. No.: 7.68e-125 Length: 3834
Score: 1237.50 Matches: 276
Percent Similarity: 38.43% Conservative: 18
Best Local Similarity: 36.08% Mismatches: 8
Query Match: 77.39% Indels: 463
Gaps: 2
DB: 3
US-10-659-004-104 (1-302) x AK028976 (1-3834)
Oy 1 MetAlaArgLysLeuSerValIleLeuIleuThrPheAlaLeuSerValThrAspPro 20
Db 500 ATGGCCACAGATTATCTGTATCATGATCTTGACCTTGCCCTTGCGTTGACAGCCCC 559
Oy 21 LeuHisGluLeuVala---AlaAlaPheProGlnThrGluLysLeuSerProAsn 39
Db 560 CTTGATACATCAACATCAACAGCGCTTCTCTGAGCTACTGAGAAATTAATCAAT 619
Oy 40 TrpGluSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTrpHisGluGln 59
Db 620 TGGAACTGATGTTAATGTGACTTGACGACTTACATGCGACGACCATCTGACGAG 679
Oy 60 LeuPheTrpArgLysGlyGluAsnAsnSerLeuSerValGluGlyPheArgLysLeu 79
Db 680 CTATTTCACCGCTACGAGAGAAATGATTCCTGTCTGTGAAGCTTCAGAAATTCCTT 739
Oy 80 GluAsnIleGlyIleAspLysIleLysArgLleHisIleHisHisAspHisHis 99
Db 740 CAGAACATAGGATAGATTAAGATTAAAGATCCATATACACCATGACACGACATCAT 799
Oy 100 SerAspHisGluHisSerAspHisGluArgHisSerAspHisGluHisSerAsp 119
Db 800 GCTACCAACGACATCACTCGACCATGAGCATCACTGACACGACATCACTCGAC 859
Oy 119 ----- 119
Db 860 CACGAGCATCACTCGACCAACGACATCACTCGACCAACGACATCCACCGCAGT 919
Oy 119 ----- 119
Db 920 CACACGTTGCTGTAATAAACAATCGAAGACCTTTGTCCAGACTTGACTGTATAT 979
Oy 119 ----- 119
Db 980 TCAGTAAAAATCCTAGAACTAGTCTAGGAGAAAGATTCGCCCAAGACAGACATGAT 1039
Oy 119 ----- 119
Db 1040 GGTAGAGAAACATCAAGAGAGTCAAGACTCTAGTGAAGTGAAGTCTGGCGGTATACAAC 1099
Oy 119 ----- 119
Db 1100 GCTGTCTCTGAAGAACTCGCTTTGTAGAGACAATAGAGACTCCAAAACCTGGAGAGCGC 1159
Oy 119 ----- 119
Db 1160 ACCAAAGATGAACCTTTTACCCCAACGACATCAACGAGAAAAGCCGAGTGGCCGG 1219
Oy 119 ----- 119

Db 1220 CTGAGTCGGCTAGCTAGGAGAGAAAGCAATGAGTCTGTGATGAGGCCAGAGAAAGCTTT 1279
 Qy 119 ----- 119
 Db 1280 ATGTATTCAGAGAAACAAATAGCAATATTGAGAGTGTTCATATCAACCAAGCTGTG 1339
 Qy 119 ----- 119
 Db 1340 ACATCCCATGGCAGTGCATCCAGGCTGTGTGAATGCACGGAATTAACCTATCTTGC 1399
 Qy 119 ----- 119
 Db 1400 CCAGCATCATCAATCAAAATTGATGCTGGGCTGTGTGATTCATACAGCAAGTAGAAG 1459
 Qy 119 ----- 119
 Db 1460 AAGCAGAAATCCCTCCAAAGACCTATTCTTACAAATAGCCTGGCTGTGTGCTTCATA 1519
 Qy 119 ----- 119
 Db 1520 GCCATTCATCATCAGTTTCTGTCTGTCTGGAGATCATCTTGTGTCCACTCATGAAC 1579
 Qy 119 ----- 119
 Db 1580 CGGGATTTTCAAGTTCCTGTGAGCTTCCTCGTGGCGCTGGCCGTGGAACGCTAGT 1639
 Qy 119 ----- 119
 Db 1640 GGGCATGCTCTGTACATCTTCTCCACACATCTCATGCAAGTCAACAGCATATAGC 1699
 Qy 119 ----- 119
 Db 1700 CATGAGAGCCAGCGATGAGAAATGAAAAGAGCCCTGTTCAGCCACTGTGCGCTCAG 1759
 Qy 119 ----- 119
 Db 1760 AATATAGAGAAAGCTCTATTTTGTGATTCAGTGTGAAGGTCTGACGGCTTAGGGGCG 1819
 Qy 119 ----- 119
 Db 1820 TTATATTTTCATGTTTCTTGTGAAACAGTACTCAGACTGATCAAGCAATTTAAAGATAG 1879
 Qy 119 ----- 119
 Db 1880 AAAAAGAGAAATCAAAAAGAACTGAAAAATGATGAGATGTGAGAGAGAGAGAGAGAGCTG 1939
 Qy 119 ----- 119
 Db 1940 TCCAAATACGACTCTCAGCTTTCCTCAATGAGAGAGTGGAGCCAGGGGAACGACT 1999
 Qy 119 ----- 119
 Db 2000 GAAAGTATCTGCGAGCCGAGCTCCAGAGAGCCCTCCCTTTGATTTCCACAGAGCCGAG 2059
 Qy 119 ----- 119
 Db 2060 ATGTGGAAGAGAGAGCTATGATGAGCCATGACACCAAGAAAGTCTACATGAA 2119
 Qy 119 ----- 119
 Db 2120 TATGTGCCAGGGGCTGCAAGAACAGTGCATTCACACTTCACAGATACGCTGGCCAG 2179
 Qy 119 ----- 119
 Db 2180 TCCGACGACCTCATCCACACCATCAGACTACATTCATTCGACACACACACACAC 2239
 Qy 120 -----HisLeuProHisSerHisSerHisLeuGlyTrpSerArgGluGluLeuLeuAspAla 137
 Db 2240 CAGAACCCACCCCTCAGACAGCAGCGGCTACTCTCAGAGAGAGCTGAAAGACGCC 2299
 Qy 138 GlyValAlaThrLeuAlaTrpMetValIleMetGlyAspGlyLeuHisAspPheSerAsp 157
 Db 2300 GGCATTCGCACATTCGGCTGTGATGATCATGGCGCAGCGGCTGCACAAATTTTCAGTAC 2359

Qy 158 GlyLeuAlaIleGlyAlaAlaPheThrGluGlyLeuSerSerGlyLeuSerVal 177
 Db 2260 GGCCTTGTATTTGGTGTGCTGCTTCCACGAGGATTTGTCCAGTGAGCTTAGAGACCTCTGTG 2419
 Qy 178 AlaValPheCysHisGlyLeuProHisGlyLeuGlyAspPheAlaValLeuLeuAla 197
 Db 2420 GCTGTGTTCTGTCAATGAATCTGCTCATGAACTAGAGTACTTGTCTGTTTCTTAAAGCT 2479
 Qy 198 GlyMetThrValIysGlnAlaValLeuValLeuTrpAsnAlaLeuSerAlaMetLeuAlaTyrLeu 217
 Db 2480 GGCATGACTGTCAAGCAGAGGCTGTGCTCTATATGCTCTGTGACATGTTGGCTTACCTT 2539
 Qy 218 GlyMetAlaThrGlyIlePheHisGlyHisTrpArgIleGluAsnValSerMetTrpIlePhe 237
 Db 2540 GGAATGGCAACAGGAGATATTCATCGGCATTATGACAGAAATGTTTCTATGTGATATTC 2599
 Qy 238 AlaLeuThrAlaGlyLeuPheMetHisValAlaLeuValAspMetValProGlyMetLeu 257
 Db 2600 GCACTCACTGCGGCTTGTTCATGATGATGCTGTGTGACATGCTGCTGAGATGTG 2659
 Qy 258 HisAsnAspAlaSerAspHisGlyCysSerArgTrpGlyTyrPhePheLeuGlnAsnAla 277
 Db 2660 CACATGATGCTAGTATGATCAGGATGACGCGTGGGATATTTCTTCTCCAGATGCT 2719
 Qy 278 GlyMetLeuLeuGlyPheGlyIleMetLeuLeuIleSerIlePheGluHisLeuVal 297
 Db 2720 GGGATACTCTCGGTTTGTGAAATGATGTACTCATTTCCATATTTGAGCATAAATGTG 2779
 Qy 298 PheArgIleAsnPro 302
 Db 2780 TTTCGTATTAATTTTC 2794
 RESULT 2
 CA488910 859 bp mRNA linear EST 14-NOV-2002
 LOCUS CA488910
 DEFINITION AGENCOURT_10808247 MAFCL Homo sapiens cDNA clone IMAGE:672115 5',
 mRNA sequence.
 ACCESSION CA488910
 VERSION CA488910.1 GI:24951701
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 859)
 AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Kristi A. Egland, Ira Pastan
 cDNA Library Preparation: Invitrogen Corp
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
 http://image.llnl.gov
 Plate: LLM14281 row: 1 column: 11
 High quality sequence stop: 755.
 FEATURES
 location/qualifiers
 1..859
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:672115"
 /cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231,
 HERT-HMB1, LNCaP"
 /lab_host="EMD108"
 /clone_id="MAFCL"
 /note="Vector: PCMV-SPORT6; Site 1: EcoRV; Site 2: Not I;
 Subtracted with brain, liver, lung, kidney and muscle.
 Directionally cloned. Priming method: oligo-dT. Average

QY 224 PheilegLyhIEtYrAlaGluAenValSerMetTrpIlePheAlaLeuThraIaGlyLeu 243
 DB 512 TTCATTGGTCATTATGCTGAAAATGTTTATATGATGATTTTGCACTTACTCTGCTTA 571
 QY 244 PheMetHsValAlaLeuValAspMetValProGluMetLeuHsAsnAspAlaSerAsp 263
 DB 572 TTCATGATAGTGTCTGCTGATATGTAACCTGAAATCTGCACATGATCTAGTAC 631
 QY 264 HsAGLyCySerAspArgTrpGlyTrpPhePheLeuGlnAsnAlaGlyMetLeuGlyPhe 283
 DB 632 CATGATATGAGCGCTGGGGGTATTTCTTTTACAGAACTGGGAGTCTTTTGAGTTT 691
 QY 284 -GlyileMetLeuLeuIleSerIlePheGluHsIleVal-PheArgIleAsnAsp 302
 DB 692 GGGATTATGTTACTTATTTCCATTTTCAACATTAATATCGTGTCTGATTAATTTTC 750
 RESULT 4 CA327621 756 bp mRNA linear EST 09-JUL-2003
 LOCUS CA327621
 DEFINITION UI-M-FY0-CCZ-1-07-0-UI.r1 NIH_BMAP_F10 Mus musculus cDNA clone
 IMAGE:6825568 5', mRNA sequence.
 CA327621
 ACCESSION CA327621.1 GI:24545719
 VERSION
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgepds-remail.nih.gov
 Tissue Procurement: Dr. Jim Loh, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)
 The following repetitive elements were found in this cDNA
 sequence: 63-144, >(TGG)nSimple_repeat (matched complement)
 Seq primer: PYX-5.
 Location/Qualifiers
 1..756
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CS7BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6825568"
 /tissue_type="whole brain"
 /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_FY0"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: Ecor I;
 Site 2: Not I; The library was constructed according
 to Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with Ecor I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is AGCGAGACAG. This library was created for the University
 of Iowa Brain Anatomy Project (BMAP). Gene Discovery in the
 Developing Mouse Nervous System, supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

ORIGIN

Alignment Scores:
 Pred. No.: 6,87e-97 Length: 756
 Score: 975.00 Matches: 190
 Percent Similarity: 92.42% Conservative: 5
 Best Local Similarity: 90.05% Mismatches: 6
 Query Match: 60.98% Indels: 11
 DB: 6 Gaps: 2
 US-10-659-004-104 (1-302) x CA327621 (1-756)
 QY 92 ILeHsHsAspHsAspHsHsSerAspHsIleGluHsHsIleSerAspHsIleGluHs 111
 DB 61 ATCCACCCACATCAGCTACCAT-----CACATTCTGCAC 96
 QY 112 SerAspHsIleGluHsHsIleSerAspHsHsIleProHsIleSerHsIleSerGlnArgTySerArg 131
 DB 97 -----CACACACACACACACACACACACCTCAGACACACACACACACGCTACTCTGA 150
 QY 132 GluGluLeuLyAspAlaGlyValAlaThrIleuAlaTrpMetValIleMetGlyAspGly 151
 DB 151 GAGGAGCTGAAGAGAGCGCGGATTCACATTCGCTGGATGTCATGATGCGGACGCG 210
 QY 152 LeuHsAspHsSerAspGlyLeuAlaIleGlyAlaAlaPheThrGluGlyLeuSerSer 171
 DB 211 CTGCACAAATTCAGTACGAGCGCTTCTATTTGCTGCTTCACCGAGGTTTGTCCAGT 270
 QY 172 GlyLeuSerThrSerValAlaValPheCyHsIleGluLeuProHsIleGluGlyAspPhe 191
 DB 271 GGGCTTAACACCTCTGCTGCTGTGTCTATGTAACCTGCCCTCAGAACTAGGTGACTTT 330
 QY 192 AlaValLeuLeuValAlaGlyMetThrValIysGlnAlaValLeuTyAsnAlaLeuSer 211
 DB 331 GCTGTTTTCGTAAGGCTGGCATGATCTGTCAAGCAGCGCTGCTCTATTAATGCTCTGTCA 390
 QY 212 AlaMetLeuAlaTyLeuGlyMetAlaThrGlyIlePheIleGlyHsIleTyAlaGluAsn 231
 DB 391 GCATGTTGGCTTACTTGGAAATGGCAACAGAGATATTCATGGGCTATTAATGCAAAAT 450
 QY 232 ValSerMetTrpIlePheAlaLeuThraIaGlyLeuPheMetHsValAlaLeuValAsp 251
 DB 451 GTTCTATATGATATTTGCACTACCTACCTGCGCTTTGTCATGATATGCTGCTTAC 510
 QY 252 MetValProGluMetLeuHsAsnAspAlaSerAspHsIleGlyCySerAspArgTrpGlyTy 271
 DB 511 ATGGTGCTGAGATGTTGCACAAATGATCTGATGTCACGATGACGCGTTGGGATAT 570
 QY 272 PhePheLeuGlnAsnAlaGlyMetLeuGlyPheGlyIleMetLeuIleSerIle 291
 DB 571 TTCTTCTGCAAGATGCTGGATATCTTCTCGTTTGGAAATTAATGTAATTCATTTCCATA 630
 QY 292 PheGluHsIleValAlaPheArgIleAsnAsp 302
 DB 631 TTGAGCATTAATTTGTCTTCTGTAATAATTTTC 662
 RESULT 5 B0541953 906 bp mRNA linear EST 13-SEP-2002
 LOCUS B0541953
 DEFINITION AGNCOURT 10325419 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6573733
 5', mRNA sequence.
 B0541953
 ACCESSION B0541953.1 GI:22852436
 VERSION
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgepds-remail.nih.gov

Db 242 ACTGTCACAGCAGCTGTGCTCTATAATGCTTTGTGACGCATGTGGCCCTACTTGGAATG 301
Qy 220 AATTTGTYLIEPHELLEGLYHISTYRALAGLUASERVASERMETTRIPLEHEALALEU 239
Db 302 GGAACAGGAGATATTCATCGGCATTAAGCAAAAATGTTCTATGTGGATATTCACACTC 361
Qy 240 TTTTATAGLYLEUPHEMETHISVALALALEUVALASPMEVALPROGLUMETLEUHSASN 259
Db 362 ACTGCCGGCTTTTCATGATATGTCGTCGTGACATGTGCTGACATGTGGACAT 421
Qy 260 AAPALASERASPHISGLCYCYSERATGTPGLTYTPHPEHELEULNASNALAGLYMET 279
Db 422 GATGCTAGTACACGAGATGACCGCTGGAGATATTTCTCTCAGAAATGCTGGAGATA 481
Qy 280 LEULGUYLPHGLYILEWETLEULEULESERILEPHEGLNHISLYSTLEVALHEARG 299
Db 482 CTTCGCGTTTGGAAATTAATGTTACTCATTTCCATTTGACCAATTAATGTTGCT 541
Qy 300 ILEASMPHE 302
Db 542 ATTAATTTTC 550

RESULT 8
LOCUS BX327149 875 bp mRNA linear EST 08-APR-2004
DEFINITION BX327149 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1023Y021 5-PRIME, mRNA sequence.
ACCESSION BX327149
VERSION BX327149.1 GI:30344608
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 875)
AUTHORS Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st enriched cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
8883.r

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0BAF0232D10_AF02189_1&c=8883.r

FEATURES

source

Location/Qualifiers
1..875
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1023Y021"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Alignment Scores:
Pred. No.: 2.22e-92 Length: 875
Score: 935.00 Matches: 191
Percent Similarity: 89.04% Conservative: 4
Best Local Similarity: 87.21% Mismatches: 22
Query Match: 58.47% Indels: 5
Gaps: 1

US-10-659-004-104 (1-302) x BX327149 (1-875)

Qy 86 LYSALLEYSAAGLIEHIEHIEHISAPRHSAPRHSISERAPRHS-----GLU 103
Db 216 AGAATTAATGCAATTCACATTTTC-CACGATACACTCGGCCAGTACAGATTCATTAC 274
Qy 104 HISISERASPHISGLUARGHISERASPHISGLUHSISERASPHISISERAPRHS 123
Db 275 CACCATATGACTACATCATATTTCTCATCATATCACACCAACCAACATCTCTAC 334
Qy 124 SERHISERGLNARGTYRSEARGGLULULEULYASAPALAGLYVALALETNLEULA 143
Db 335 AGTACACGCCAGCCCTACTCTCGGAGAGAGCTGAAATGCGCGCTGCGCATCTGGCC 394
Qy 144 TTPMETVALLIEMETGLYASPGLYLEUHSAPRHSERASPHISGLYVALALETNLEULA 163
Db 395 TGGATGGGATTAAGGGGATGAGGCTGCACATTTTCAGCGATGCGCTGACATTTGCT 454
Qy 164 AIAPEHTRGUGLYLEUSERSERGLYLEUSERTHSERVALAIAVALPHECYANISGLU 183
Db 455 GCTTTTACTGAGAGCTTATCATAGTGTTTAAGTACTTCTGCTGCTTCTGTCATGAG 514
Qy 184 LEUPROHISGLULEUGLYASPPHEALIAVALLEULEULYBALAGLYMETTHVALISGLIN 203
Db 515 TTGCTCATGATTAAGTGAAGTCTTGCTGTTCTACTAAAGCTGCGCATGACCGTTAAGCAG 574
Qy 204 AIAVALLEULYRANALALEUSERALAMETLEULALATYLEUGLYMETALATHRGLYLE 223
Db 575 GCTGCTCTTAAATGATGATGTCAGCCATGCTGCGATGCTTGAATGGCAACAGAAAT 634

Qy 224 PHEILEGLYHISTYRALAGLUASNVASERMETTRIPLEHEALALEUThRALAGLYLEU 243
Db 635 TTCAATGGTCAATTATGCTGAAGAAATCTTTCTATGTGATTTGCACTTACTGCTGGCTTA 694
Qy 244 PHEMETHISVALALALEUVALASPMEVALPROGLUMETLEUHSAPRHSAPRHSERAP 263
Db 695 TTCAATGATGTTGCTGCTGTTGATATGCACTTGAATGCGCAATGATGATGATGAC 754
Qy 264 HISGLCYCYSERATGTPGLTYTPHPEHELEULNASNALAGLYMETLEULEULYPHE 283
Db 755 CATGATGATGACCCCTGGGGTATTTCTTTTACAGAT-GCTGGAGATGCTTTGGGTTTT 813
Qy 284 GLYILEWETLEULEULESERILEPHEGLNHISLYSTLEVALHEARGILEASMPHE 302
Db 814 GGAAT-ATGTTACTTATTTCCATATTTGACATTAATAATCGGGTTTCGATATATTTTC 869

RESULT 9
LOCUS CA454784 893 bp mRNA linear EST 12-NOV-2002
DEFINITION AGENCOURT 10763039 MABCL Homo sapiens cDNA clone IMAGE:6721427 5',
mRNA sequence.

ACCESSION CA454784
VERSION CA454784.1 GI:24904853
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 893)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov

Tissue Procurement: Kristi A. Eglund, Ira Pastan
cDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
http://image.llnl.gov
Plate: LHAM14282 Row: 1 Column: 11
High quality sequence stop: 593.

FEATURES
SOURCE

Location/Qualifiers
1..893
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6721427"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HME1, LNCaP"
/lab_host="EMD110B"
/clone_lib="MAPCL"
/note="Vector: pCMV-SPORT6; Site 1: EcorV; Site 2: Not I; Subcloned with brain, liver, lung, kidney and muscle. Directionally cloned. Priming method: oligo-dT. Average insert size: 1800 bp. Library amplification: 26,000 fold. Kristi A. Egland, James J. Vincent, Robert Strausberg, Bunkook Lee & Ira Pastan: Discovery of new breast cancer genes encoding membrane and secreted proteins. Manuscript submitted."

ORIGIN

Alignment Scores:
Pred. No.: 3.65e-89 Length: 893
Score: 906.00 Matches: 175
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.43% Mismatches: 0
Query Match: 56.66% Indels: 0
DB: 6 Gaps: 0

US-10-659-004-104 (1-302) x CA454784 (1-893)

QY 127 GlnAgtTyrSerArgGluGluLeuYsaAplAgiValAlaThrLeuAlaTrpMetVal 146
Db 2 CAGGCGTACTCTCGGAGAGCTGAAGAAGTCCGGCGTCCACCTCGCCGTGAGTGG 61

QY 147 IlleMetGlyAspGlyLeuHisAsnPheserApgGlyLeuAlaIleGlyAlaAlaPheThr 166
Db 62 ATAAATGGGTGATGGCTGCACAAATTCAGCCATGGCCCTACGAAATGGCTGCTTACT 121

QY 167 GluGlyLeuSerSergGlyLeuSerThrSerValAlaValPheCysHisGluLeuProHis 186
Db 122 GAAGCTTATCAAGTGGTTAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 181

QY 187 GluLeuGlyAspPheAlaValLeuLeuYsaAplAgiMetThrValGlyAlaValLeu 206
Db 182 GAATTAAGGTGACTTGGTCTTACTTAAAGCTGGCAATGACCGTTAAACAGCTGCTCTT 241

QY 207 TyrAsnAlaLeuSerAlaMetLeuAlaTyrLeuGlyMetAlaThrGlyIlePheIleGly 226
Db 242 TATAATGCAATGTCAGCCATGCTGCGCTATCTTGGAAATGGCAACGAAATTTCTTGGT 301

QY 227 HisTyrAlaGluAsnValSerMetTyrIlePheAlaLeuThrAlaGlyLeuPheMetHis 246
Db 302 CATTAATGCTGAATAATGTTCTATGATGATATTGCACTTACGCTGCTGCTTATTCATGAT 361

QY 247 ValAlaLeuValAspMetValProGluMetLeuHisAsnAplAserAspHisGlyCys 266
Db 362 GTTGCTGCTGTTGATATGTAATGCTGAATGCTGCAATGATGTAAGTACATGAGATG 421

QY 267 SerArgTyrGlyTyrPhePheLeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMet 286
Db 422 AGCCGCTGGGGGATTTCTTTTACAGAAATGCTGGAGATGCTTTGGGTTTGGAAATTTG 481

QY 287 LeuLeuIleSerIlePheGlnHisIleValPheArgIleAsnPhe 302
Db 482 TTACTTATTTCCATATTGGAACATTAATCGTGTTCTGTAATAATTTTC 529

RESULT 10
BP259977 582 bp mRNA linear EST 16-SEP-2004
LOCUS BP259977 Sugano cDNA library, heart Homo sapiens cDNA clone
DEFINITION HRT08420, mRNA sequence.
ACCESSION BP259977
VERSION BP259977.1 GI:52175207

KEYWORDS
SOURCE EST.
ORGANISM Homo sapiens (human)
REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 582)
AUTHORS Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.

FEATURES
SOURCE

Location/Qualifiers
1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HRT08420"
/tissue_type="heart"
/clone_lib="Sugano cDNA library, heart"

ORIGIN

Alignment Scores:
Pred. No.: 5.31e-88 Length: 582
Score: 893.00 Matches: 177
Percent Similarity: 99.44% Conservative: 1
Best Local Similarity: 98.88% Mismatches: 1
Query Match: 55.85% Indels: 1
DB: 5 Gaps: 0

US-10-659-004-104 (1-302) x BP259977 (1-582)

QY 124 SerHisSerGlnAgtTyrSerArgGluGluLeuYsaAplAgiValAlaThrLeuAla 143
Db 2 AGTACACAGCCAGCTACTCTCGGAGAGCTGAAGA-GCCGGCGTCCACCTCTGACC 60

QY 144 TrpMetValIleMetGlyAspGlyLeuHisAsnPheserApgGlyLeuAlaIleGlyAla 163
Db 61 TGGATGGTGAATAGGGTGAATGGCTGCACAAATTCAGCCATGGCTGCAATGGTCTCT 120

QY 164 AlaPheThrGluGlyLeuSerSergGlyLeuSerThrSerValAlaValPheCysHisGlu 183
Db 121 GCTTTTACTGAAGCTTATCAAGTGGTTAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180

QY 184 LeuProHisGluLeuGlyAspPheAlaValLeuLeuYsaAplAgiMetThrValIleGln 203
Db 181 TTGCTCATGAATTAAGGTGACTTGGTCTTACTTAAAGCTGGCAATGACCGTTAAAGCAG 240

QY 204 AlaValLeuTyrAsnAlaLeuSerAlaMetLeuAlaTyrLeuGlyMetAlaThrGlyIle 223
Db 241 GCTGCTCTTAAATGCAATTCAGCCATGCTGCGCTATCTTGGAAATGGCAACGAAATTT 300

QY 224 PheIleGlyHisTyrAlaGluAsnValSerMetTyrIlePheAlaLeuThrAlaGlyLeu 243
Db 301 TTCAATGGCTATTAATGCTGAATAATGTTCTATGATGATATTGCACTTAATGCTGCTGCTTA 360

QY 244 PheMetHisValAlaLeuValAspMetValProGluMetLeuHisAsnAplAserAsp 263
Db 361 TTCATGATGTGCTGCTGTTGATGATGCTGAATGCTGCAATGATGCTGATGAC 420

QY 264 HisGlyCysSerArgTyrGlyTyrPhePheLeuGlnAsnAlaGlyMetLeuLeuGlyPhe 283
Db 421 CATGATGTACCCGCTGGGGATTTCTTTTACAGAAATGCTGGAGATGCTTTGGGTTT 480

QY 284 GlyIleMetLeuLeuIleSerIlePheGlnHisIleValPheArgIleAsnPhe 302
Db 481 GGAATTAATGTAATTTATTTCCATATTGGAACATTAATCGTGTTCTGTAATAATTTTC 537

RESULT 11

ORIGIN

C, or G and N = A, C, G, or T). Average insert size 1.9 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH-MGC Library."

Alignment Scores:

Pred. No.:	8,1e-86	Length:	791
Score:	875.00	Matches:	169
Percent Similarity:	99.42%	Conservative:	2
Best Local Similarity:	98.26%	Mismatches:	1
Query Match:	54.72%	Indels:	0
DB:	4	Gaps:	0

US-10-659-004-104 (1-302) x BG547794 (1-791)

```

Qy 131 ArggluGlueuLysAspAlaGlyValAlaThrLeuAlaTrpMetValIleMetGlyAsp 150
Db 1 CGGAGAGAGCTGAAGAAGATGCCGGCGTCCCACTTGGCTGGATGGTATATGGGTAT 60
Qy 151 GlyLeuHisAsnProSerAspGlyLeuAlaIleGlyAlaAlaPheThrGlyLeuSer 170
Db 61 GGCGTCACATTTTCAGCGATGCCCTACCAATGGTCTCTTTACTGAAGGCTTATCA 120
Qy 171 SerGlyLeuSerThrSerValAlaValPheCysHisGlyLeuProHisGlyLeuGlyAsp 190
Db 121 AGTGGTTTAAGTACTTCTGTTGCTGTTCTGTCATGAGTGTCCTCCTCAAGAAATAGGTAC 180
Qy 191 PheAlaValLeuLeuLysAlaGlyMetThrValLysGlnAlaValLeuTyrAsnAlaLeu 210
Db 181 TTGCTGTTCTACTTAAGAGCTGGCATGACCGTTAAGCAGGCTGCTTATTAAGCATG 240
Qy 211 SerAlaMetLeuAlaTyrLeuGlyMetAlaThrGlyIlePheIleGlyHisTyrAlaGlu 230
Db 241 TCAGCCATGCTGGCGATCTTGGAAATGCAACAGAAATTTTCATTGCTCATTAATGCGAA 300
Qy 231 AsnValSerMetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisValAlaLeuVal 250
Db 301 AATGTTTCTATGAGATATTTGCACTTACTGCTGCTTATTCATGATGATGTTGCTGTT 360
Qy 251 AspMetValProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSerArgTrpGly 270
Db 361 GATATGCACTGAAGATGCTGCACATATGCTAGTACCATGATGATGAGCGCGTGGGG 420
Qy 271 TyrPheLeuGlnAsnAlaGlyMetLeuGlyPheGlyIleMetLeuLeuIleSer 290
Db 421 TATTTCTTTTACAGAAATGCTGGATGCTTATGGGTTCTGGAATTAATGTTACTTATTTCC 480
Qy 291 IlePheGluHisValIleValPheArgIleAsnPro 302
Db 481 ATATTGAAACATAAATCGTGTTCGTATTAATTTTC 516

```

RESULT 13
CB052402
LOCUS 539 bp mRNA linear EST 17-JAN-2003
DEFINITION NISC_S107906.y1 NCI_CGAP_Le12 Homo sapiens cDNA clone IMAGE:3290314
5', mRNA sequence.
ACCESSION CB052402
VERSION CB052402.1 GI:27790689
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 539)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
Tumor Gene Index
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapdb-remail.nih.gov
cDNA Library Preparation:

FEATURES

SOURCE

cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LNL
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC)
Sequencing distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
info@image.lnl.gov
Plate: L1AM8057 row: N column: 11
Seq primer: M13RP1 reverse primer (ABI).
Location/Qualifiers

ORIGIN

Alignment Scores:

Pred. No.:	2.66e-83	Length:	539
Score:	850.00	Matches:	162
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	97.01%	Mismatches:	5
Query Match:	53.16%	Indels:	0
DB:	6	Gaps:	0

US-10-659-004-104 (1-302) x CB052402 (1-539)

```

Qy 136 AspAlaGlyValAlaThrLeuAlaTrpMetValIleMetGlyAspGlyLeuHisAsnPro 155
Db 21 GAGCCGCGCGCTCTCACCTTGGCTGATGCTGATTAATGGGTGATGGCTGCACAAATTC 80
Qy 156 SerAspGlyLeuAlaIleGlyAlaAlaPheThrGlyLeuSerSerGlyLeuSerThr 175
Db 81 AGCGATGCGCTTAAGAAATGGTGTCTTTTACTGAAGCTTATCAAGGTGTTAAGTACT 140
Qy 176 SerValAlaValPheCysHisGlyLeuProHisGlyLeuGlyAspPheAlaValLeuLeu 195
Db 141 TCTGTTGCTGTTCTGTCATGAGTTCCTCATATAATTAAGTGACTTTTCTGTTCTACTA 200
Qy 196 LysAlaGlyMetThrValLysGlnAlaValLeuTyrAsnAlaLeuSerAlaMetLeuAla 215
Db 201 AAGGCTGACATGACCGTTAAGCAGGCTGTCTTATATGATGATGTCAGCATCTGGCG 260
Qy 216 TyrLeuGlyMetAlaThrGlyIlePheIleGlyHisTyrAlaGluAsnValSerMetTrp 235
Db 261 TATCTTGGAAATGGGAAGAAGATTTTCAATGATGATTAATGAGAAATGTTTCTATGGG 320
Qy 236 IlePheAlaLeuThrAlaGlyLeuPheMetHisValAlaLeuValAspMetValProGlu 255
Db 321 ATATTTCACACTTACTGCTGCTTATTCATGATATGCTCTCGTGTGATATGTAACCTGAA 380
Qy 256 MetLeuHisAsnAspAlaSerAspHisGlyCysSerArgTrpGlyTyrPhePheLeuGln 275
Db 381 ATGCTGACACATGATGCTAGTACCATGATGATGATGATGATGATGATGATGATGATGAT 440
Qy 276 AsnAlaGlyMetLeuGlyPheGlyIleMetLeuLeuIleSerIlePheGluHisLys 295
Db 441 AATGCTGGAGATCTTTGGGTTTGGAAATTAATGTTACTTATTTCCATTAATTAATTAATA 500
Qy 296 IleValPheArgIleAsnPro 302
Db 501 ATCGTGTTCGTGTAATTTTC 521

```



```

/dev stage="Adult"
/clone_lib="Brl33"
/note="Organ: breast; Vector: puc18; Site_1: Sma1; Site_2:
Sma1; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

```

ORIGIN

Alignment Scores:

Pred. No.:	1.8e-82	743
Score:	846.00	Matches: 181
Percent Similarity:	86.05%	Conservative: 4
Best Local Similarity:	84.19%	Mismatches: 13
Query Match:	52.91%	Indels: 17
DB:	1	Gaps: 2

US-10-659-004-104 (1-302) X AI907176 (1-743)

[illegible]

Search completed: March 20, 2005, 01:36:15
Job time : 3705 secs

This Page Blank (uspto)